European Psychiatry

www.cambridge.org/epa

Impact of early risk factors on schizophrenia risk and age of diagnosis: A Danish population-based register study

Cecilie K. Lemvigh[1], Karen S. Ambrosen[1], Bjørn H. Ebdrup[1,2], Birte Y. Glenthøj[1,2], Merete Osler[3,4] and Birgitte Fagerlund[5,6]

[1]Center for Neuropsychiatric Schizophrenia Research (CNSR), Mental Health Center, Glostrup, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark; [2]Faculty of Health and Medical Sciences, Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; [3]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospitals, Copenhagen, Denmark; [4]Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark; [5]Child and Adolescent Mental Health Center, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark and [6]Faculty of Social Sciences, Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Research Article

Cite this article: Lemvigh CK, Ambrosen KS, Ebdrup BH, Glenthøj BY, Osler M, Fagerlund B (2024). Impact of early risk factors on schizophrenia risk and age of diagnosis: A Danish population-based register study. *European Psychiatry*, **67**(1), e64, 1–8
https://doi.org/10.1192/j.eurpsy.2024.1774

Received: 26 February 2024
Revised: 19 June 2024
Accepted: 10 July 2024

Keywords:
age of onset; child-onset; early-onset; psychosis; risk stratification

Corresponding author:
Cecilie K. Lemvigh;
Email: cecilie.koldbaek.lemvigh@regionh.dk

Abstract

Background. While several risk factors for schizophrenia have been identified, their individual impacts are rather small. The relative independent and cumulative impacts of multiple risk factors on disease risk and age of onset warrant further investigation.

Study design. We conducted a register-based case–control study including all individuals receiving a schizophrenia spectrum disorder in Denmark from 1973 to 2018 (N = 29,142), and a healthy control sample matched 5:1 on age, sex, and parental socioeconomic status (N = 136,387). Register data included parental history of psychiatric illness, birth weight, gestational age, season of birth, population density of birthplace, immigration, paternal age, and Apgar scores. Data were analysed using logistic regression and machine learning.

Results. Parental history of psychiatric illness (OR = 2.32 [95%CI 2.21–2.43]), high paternal age (OR = 1.30 [1.16–1.45]), and low birth weight (OR = 1.28 [1.16–1.41]) increased the odds of belonging to the patient group. In contrast, being a second-generation immigrant (OR = 0.65 [0.61–0.69]) and high population density of the birthplace (OR = 0.92 [0.89–0.96]) decreased the odds. The findings were supported by a decision tree analysis where parental history, paternal age, and birth weight contributed most to diagnostic classification (ACC_{test} = 0.69, AUC_{test} = 0.59, $p < 0.001$). Twenty percent of patients were child-onset cases. Here, female sex (OR = 1.82 [1.69–1.97]) and parental psychiatric illness (OR = 1.62 [1.49–1.77]) increased the odds of receiving the diagnosis <18 years.

Conclusion. Multiple early factors contribute independently to a higher psychosis risk, suggesting cumulative effects leading to symptom onset. Routine assessments of the most influential risk factors could be incorporated into clinical practise. Being female increased the risk of diagnosis during childhood, suggesting sex differences in the developmental trajectories of the disorder.

Introduction

Schizophrenia is hypothesised to be a neurodevelopmental disorder emerging many years prior to manifestation of overt psychotic symptoms and when the clinical diagnosis is made [1]. A neurodevelopmental condition is conceptualised as starting before birth with genes and environmental factors interacting during different phases of development, resulting in a neurobiological vulnerability and premorbid impairments in neuromotor abilities, language and social skills, and cognitive development [2]. Early identification of individuals at risk for schizophrenia and establishment of preventive strategies are critical steps towards improving the outcome for this patient group. Here, one major challenge is the development of reliable and efficient means to predict psychosis [3].

Although schizophrenia is a highly heritable disorder [4], genetic factors alone only partially account for all clinical cases as reflected by the 50% disease discordance rate in monozygotic twin pairs [4, 5]. Decades of research have yielded substantial evidence for the importance of early risk factors in the development of schizophrenia. These include low birth weight, high paternal age, birth complications, immigration, urbanicity, and winter/spring birth [6–8]. Previous register studies, including Danish cohorts, have contributed substantially to this line of research. For example, using Danish register data, a family history of mental illness [9], low birth weight and premature birth [10], high paternal age (with a stronger association for female offspring) [11], urbanicity [9], immigration [12], substance use disorder (highest risk for cannabis or alcohol)

[13], and male sex [14] have been shown to increase the risk of schizophrenia. Here, it is important to note that exposure to these factors is relatively common, while the prevalence of schizophrenia in comparison is low [15]. Moreover, the observed effects are often small [8] and consequently, the explanatory power of any risk factor alone is quite low. Thus, it is likely that a genetic vulnerability interacting with multiple risk factors cumulating over time leads to onset of psychotic symptoms once a certain threshold has been reached [16, 17].

Fewer studies have investigated more than two risk factors simultaneously [18–20]. One study created a so-called polyenviromic risk score by combining multiple known environmental risk factors, each weighted by its odds ratio for the association with schizophrenia in the literature [19]. In this study, a higher polyenviromic risk score was significantly associated with conversion to psychosis in young individuals at familial high risk for psychosis. Another study showed an association between the number of environmental risk exposures and age of onset for patients with schizophrenia. Here, individuals exposed to four or more risk factors received the diagnosis nearly nine years earlier than those with no known exposure [20].

The typical age of psychotic symptom onset is late adolescence to early adulthood [21], with a peak age of symptom onset in the early twenties [22]. An early onset defined as diagnosis before the age of 18 has been associated with a poorer prognosis [23, 24], higher rates of substance abuse [25], lower educational achievement [25], and a stronger familial disposition for schizophrenia that may be more pronounced for females compared to males [26]. Diagnosis before age 13 is considered very early onset and has been associated with more severe premorbid neurodevelopmental disturbances and a greater familial vulnerability [2].

The majority of the abovementioned risk factors are established several years before symptom onset, making them susceptible to recall bias if patients are interviewed at the time of diagnosis. The use of register data obtained at birth provides one strategy to include prospective information on early risk factors in a case–control study [18, 28]. The current study provides an opportunity to weigh the relative importance of each risk factor against each other in a large and highly representative sample. This study includes all individuals who received a schizophrenia spectrum diagnosis in Denmark during the time period, covering both child- and adult-onset cases.

The overall aim of the study is to examine both the independent and cumulative or interactive impact of multiple risk factors on the development of schizophrenia spectrum disorders in a population-based case–control register study. As a secondary aim, we investigate how the included risk factors influence the age of diagnosis. Based on the previous literature, we hypothesise: 1) multiple risk factors will increase the risk of schizophrenia spectrum disorders, either cumulatively or interactively, 2) the early risk factors will increase the risk of diagnosis during childhood or adolescence, and 3) a family history of psychiatric illness (as a proxy measure for the genetic impact) will show the strongest association with schizophrenia risk/age of diagnosis and show interactions with other early risk factors. As an exploratory aim, we also examine potential sex differences.

Methods

The study was approved by the regional Data Inspection. No individual-level consent was required, and all data used were pseudo-anonymized (permission number: P-2020-88, Statistics Denmark permission: 707913).

Study population

In this nationwide register-based case–control study, we included all individuals in Denmark with a first ICD-8 or ICD-10[1] schizophrenia spectrum diagnosis registered in the Danish National Patient Registry from 1973 through 2018. The reference population included five healthy controls (HCs) per patient and was selected from the general Danish population, who did not have any psychiatric hospital contacts using coarsened exact matching with age, sex, and parental socioeconomic status (labour market affiliation, income, and years of education) as match variables. The dataset consisted of 165,529 individuals, including 29,142 patients with schizophrenia spectrum disorders and 136,387 HCs. Age of first diagnosis <18 years old was considered as child−/early onset cases and < 13 years as very early onset. Evidence suggest that schizophrenia diagnoses are valid in children as young as 7 years old [23] and thus cases with an age of diagnosis <7 years old were excluded (N = 32).

Risk factors

Information on early life risk factors were collected from several Danish registers [29], including the medical birth registry [30], the Danish Psychiatric Central Research Register [31], and The National Patient Register [32]. The Civil Personal Register number assigned to all persons in Denmark allowed for individual-level record linkage between the registers [33]. We included the following risk factor variables [8]: parental history of psychiatric illness (as a proxy for genetic influences), birth weight (low <2500 g and ultra-low <2000 g) [6, 28], gestational age (premature birth <36 weeks), winter/spring birth (December–May) [34], population density of birthplace (as a measure of urbanicity) [9, 27], paternal age (high >45 years) [35], and immigration status (Danish origin or second-generation immigrant) [36]. Because the focus of this study was on data from the medical birth registry, only individuals born in Denmark could be included and therefore the sample does not include any first-generation immigrants. From this register, we also included Apgar scores [37], which is a quick test performed routinely on the newborn one and five minutes after birth, and includes assessment of activity (muscle tone), heart rate, grimace (reflex irritability), appearance (skin colour), and breathing. Each category is scored from 0 to 2, with higher scores indicating better functioning. Due to a higher degree of missing data for Apgar scores after one minute, we only included the Apgar score obtained after five minutes in the analyses both as a continuous measure and based on the following cut-off values (a score of 7–10 was defined as normal, 4–6 as abnormal, and a score of 0–3 as low) [37].

Statistical analyses

Statistical analyses were performed using SPSS (version 27.0, SPSS Inc.), R (version 4.1.0), and Python (version 3.7). First, all variables were examined for outliers, and implausible scores were removed (e.g., Apgar scores >10). We included the variance inflation factor (VIF) to examine multicollinearity between

[1]**ICD-10 codes**: F20 schizophrenia, F21 schizotypal disorder, F22 delusional disorders, F23 brief psychotic disorder, F24 shared psychotic disorder, F25 schizoaffective disorders, F28 other psychotic disorder not due to a substance or known physiological condition, F29 unspecified psychosis not due to a substance or known physiological condition. **ICD-8 codes**: 295 schizophrenia, 297 paranoid states, 298 other psychosis, 299 unspecified psychosis, 301.09 personality disorder paranoid 301.29 personality disorder schizoid.

https://doi.org/10.1192/j.eurpsy.2024.1774 Published online by Cambridge University Press

the included independent variables, and continuous variables were demeaned to remove collinearity. All VIF values were between 1.00 and 2.66.

Logistic regression was applied because of the interpretability of odds ratios to examine the relative importance of the included risk factors (patients vs HC), and for age of diagnosis (child- vs adult-onset cases). Because our case–control study was based on loose matching data on a few demographic variables, we chose to run unconditional logistic regression models [38]. We tested for multiplicative interactions between all significant predictors. In addition, because of previous literature suggesting gene–environment interactions in schizophrenia, we also decided to run interactions between parental history of psychiatric illness and all other risk factors. For patients vs HCs, we included risk factors as both continuous variables and categorical based on cut-off values presented in the literature. The amount of missing data was compared between groups, and as a sensitivity analysis, we also ran the model with two types of imputation, that is, median and multiple imputation by chained equations (MICE) using the *IterativeImputer* function in scikit-learn [39]. As a post hoc analysis, we reran the model of child- versus adult-onset cases, limiting the child group to patients with a very early onset (diagnosis <13 years).

To examine the predictive power of the early risk factors, we adopted a gradient boosting classifier with decision trees as week learners, as implemented in the Extreme Gradient Boosting (XGBoost, version 0.90) open-source software library [40] (https://xgboost.ai/). We used SHapley Additive exPlanations (SHAP) plots to evaluate how much each feature affected the prediction as an estimate of the relative contribution of each risk factor variable. The XGBoost classifier was trained using stratified 10-fold cross-validation. All parameters had default values, except the '*scale_pos_weight*' that were set to the ratio between the two classes, such that the samples in the minor class were given higher weights. P-values were obtained by permutation tests with 1,000 permutations as implemented in the scikit-learn library (version 0.21.3) [39]. The predictive performance by the XGBoost classifier was assessed using accuracy and area under the receiver operating characteristic curve on the test set (ACC_{test} and AUC_{test}, respectively). Because of a very skewed amount of missing data between the child and adult patient groups, Apgar scores, paternal age and gestational age were removed from the decision tree analyses.

As another sensitivity analysis, we explored the effects of time by analysing the data separately for four decades (1970s, 1980s, 1990s and 2000s).

Results

Average demographic and risk factor information for HCs and patients (combined and divided on child- and adult-onset cases) are presented in Table 1. As expected, there were no differences in age or sex between HCs and patients. The mean age of the first schizophrenia spectrum diagnosis for patients was ~23 years old (see Supplementary Table S1 for diagnosis distribution).

Table 1. Demography and risk factor variables for HCs and patients with schizophrenia (combined and split on age of diagnosis)

	HCs (*N* = 136,387)	Patients (*N* = 29,142)	Adult-onset (*N* = 23,683)	Child-onset (*N* = 5,427)
Sex (female/male), %	44%/56%	44%/56%	41%/59%	56%/44%
Age, Mean (SD)	34.59 (8.10)	34.51 (8.08)	36.14 (7.35)	27.21 (7.21)
Age of diagnosis, Mean (SD)	–	22.81 (6.4)	24.61 (5.66)	14.96 (2.07)
Paternal age, Mean (SD)	30.72 (5.65)	31.01 (6.05)	30.82 (5.98)	31.67 (6.25)
High paternal age > 45 (yes/no), %	2%/98%	3%/97%	3%/97%	3%97
Birth weight (g), Mean (SD)	3,362.28 (585.05)	3,303.59 (623.66)	3,286.94 (622.50)	3,370.88 (624.62)
Low birth weight < 2,500 g (yes/no), %	5%/95%	7%/93%	7%/93%	7%/93%
Ultra-low birth weight < 2,000 g (yes/no), %	2%/98%	3%/97%	3%/97%	3%/97%
Gestational age (weeks), Mean (SD)	39.53 (1.83)	39.41 (2.04)	39.43 (2.03)	39.35 (2.07)
Premature birth <36 weeks (yes/no), %	3%/97%	4%/96%	4%/96%	4%/96%
Apgar score (after 5 min), Mean (SD)	9.86 (0.71)	9.85 (0.76)	9.85 (0.75)	9.85 (0.78)
Abnormal Apgar <7 (yes/no), %	1%/99%	1%/99%	1%/99%	1%/99%
Low Apgar 0–3 (yes/no), %	0.3%/99.7%	0.3%/99.7%	0.3%/99.7%	0.4%/99.6%
Winter birth (yes/no), %	50%/50%	50%/50%	50%/50%	49%/51%
Immigration status, %				
Ethnic Danish	93%	95%	95%	94%
Second-generation immigrants	7%	5%	5%	6%
Population density of birthplace, %				
> 100,000	31.6%	30.1%	29.5%	33.2%
10,000–100,000	68.3%	69.7%	70.3%	66.7%
< 10,000	0.1%	0.2%	0.2%	0.1%
Parental history of psychiatric diagnosis (yes/no), %	10%/90%	21%/79%	19%/81%	27%/73%

https://doi.org/10.1192/j.eurpsy.2024.1774 Published online by Cambridge University Press

Patients versus HC

Figure 1A shows the results from logistic regression examining the risk of psychosis. Parental history of psychiatric illness (Odds ratio (OR) = 2.32, 95% CI [2.21–2.43]), high paternal age (OR = 1.30 [1.16–1.45]), and low birth weight (OR = 1.28 [1.16–1.41]) significantly increased the risk of belonging to the patient group. In contrast, being a second-generation immigrant (OR = 0.65, [0.61–0.69]), and high population density of the birthplace (OR = 0.92 [0.89–0.96]) significantly decreased the risk.

Rerunning the model using continuous variables instead of categorical ones showed similar results, although with slightly lower odds ratios (Supplementary Table S2). Including interactions in the model only revealed a significant negative interaction between parental history of psychiatric illness and immigration status (OR = 0.77 [0.65–0.92]). All other interactions were not significant (all p-values >0.09). Sex was not significant in the model and rerunning the analysis separately for females and males showed similar influences of the included risk factors across sexes (Supplementary Table S2).

These findings were in line with results from the XGBoost classifier, where parental history of psychiatric illness, paternal age and birth weight contributed most to the classification (average ACC_{test} = 0.70 (SD = 0.006), p < .001, average AUC_{test} = 0.59 (SD = 0.007), p < .001) (Figure 2A).

Child versus adult onset

Approximately 20% of patients (N = 5,427) were classified as child-onset cases (diagnosis before <18 years). The mean age of diagnosis was 25 years for the adult-onset cases and 15 years for the child-onset cases. Female sex (OR = 1.82 [1.69–1.97]), and parental history of psychiatric illness (OR = 1.62 [1.49–1.77]) increased the risk of a diagnosis before turning 18 years old with the highest odds ratios. Higher birth weight (OR = 1.02 [1.01–

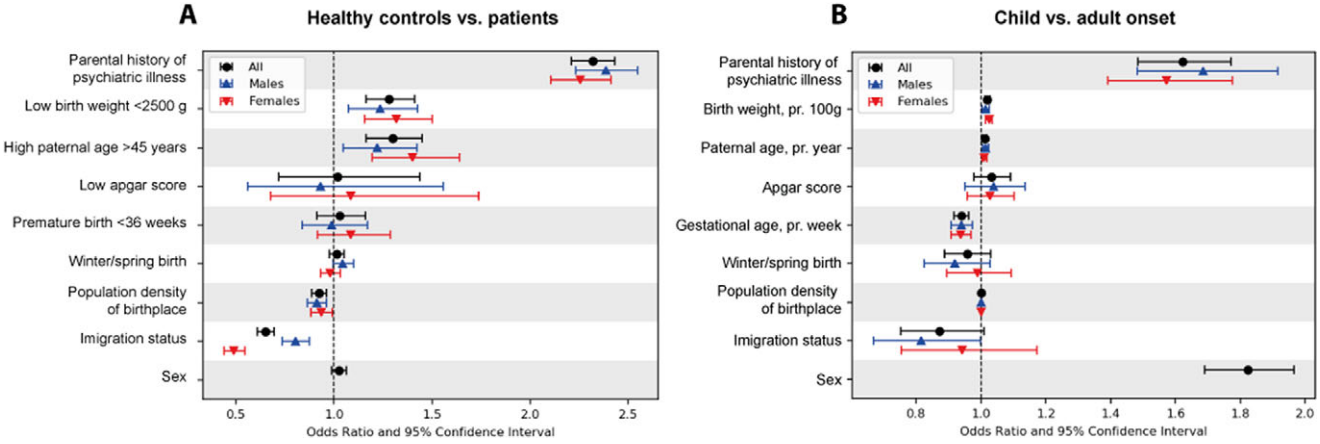

Figure 1. Forrest plots of early risk factors. Legend: Forrest plots showing odds ratios from logistic regression models of A) healthy controls versus patients, and B) within the patient group between child-onset and adult-onset cases. Adjusting for age did not change the findings (estimates not shown).

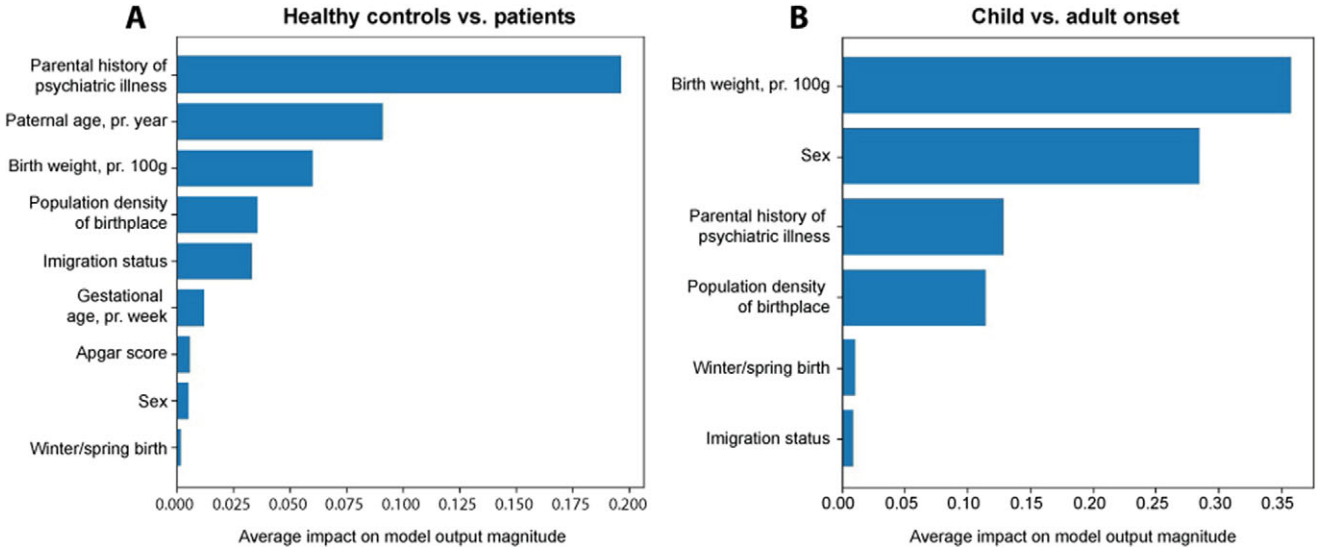

Figure 2. SHAP plots of early risk factors. Legend: SHAP plots from decision tree analyses showing the relative importance of each variable for the classification. A) healthy controls versus patients, and B) within the patient group between child-onset and adult-onset cases. Because of a very skewed amount of missing data between child vs adult patient groups, Apgar scores, paternal age and gestational age were removed from the decision tree analyses resulting in fewer variables in panel B.

https://doi.org/10.1192/j.eurpsy.2024.1774 Published online by Cambridge University Press

1.03]), higher paternal age (OR = 1.01, [1.01–1.02]), and higher population density of birthplace (OR = 1.00, [1.00–1.00]) also significantly increased the risk of onset during childhood, while longer gestational age (OR = 0.94 [0.92–0.96]) decreased the risk (Figure 1B and Supplementary Table S3).

We did not find any significant interactions (all p-values >0.11) nor any sex differences when rerunning exploratory analyses split on sex (Supplementary Table S3).

Results from the decision tree analysis showed that birth weight and sex contributed most to the classification (average ACC_{test} = 0.61 (SD = 0.010), $p < 0.001$, average AUC_{test} = 0.67 (SD = 0.013), $p < 0.001$) (Figure 2B).

As a post hoc analysis, we focused on the group of patients characterized as very early onset cases. Approximately 2% of the patients (N = 586) were diagnosed before the age of 13 years. Rerunning the analyses comparing this group to adult-onset cases revealed a similar pattern of findings, although the previously observed effect for the female sex was no longer significant in this smaller group. In contrast, Apgar scores and immigration status here reached significance (Supplementary Material Table S4 and Figure S1).

Sensitivity analyses

The amount of missing data is presented in Supplementary Table S5. Imputation of missing data using either median or MICE imputation yielded very similar findings (Supplementary Table S6). Splitting the data into four decades revealed a smaller sample size for the early and late time periods, respectively (Supplementary Table S7). Overall, similar findings were observed across decades, indicating that the findings were not driven by a specific time period. In the smaller sample of individuals born between 1973 and 1979, only parental history of psychiatric illness reached significance, however, overall, the odds ratios were very consistent across time.

Discussion

In this population-based case–control study, we showed that parental history of psychiatric illness, high paternal age, and low birth weight independently increased the risk of developing a schizophrenia spectrum disorder. We observed no interactions between the significant risk factors, suggesting a cumulative impact. Moreover, parental history of psychiatric illness and female sex increased the risk of receiving the diagnosis early (<18 years). The study included all individuals born in Denmark, who received a first diagnosis of a schizophrenia spectrum disorder between 1973 and 2018 and a matched HC sample.

In line with our first hypothesis, several factors were independently associated with the risk of developing a schizophrenia spectrum disorder. Parental history of psychiatric illness had the strongest impact on our models. Moreover, low birth weight (<2500 g), and high paternal age (>45 years) also significantly increased the risk in line with the previous literature [6, 10, 11]. However, the observed odds ratios were lower compared to previous studies, which may be explained by the fact that we adjusted for all the other risk factors included in the model. We did not find any significant influences of winter/spring birth, low Apgar scores or premature birth. The finding of no effect of season of birth is in line with some previous findings from Denmark [9].

Surprisingly, being a second-generation immigrant and having a high population density of the birthplace actually decreased the risk of belonging to the patient group in this cohort. These results are in contrast with previous evidence [36, 41], also from Danish cohorts [9, 12].

For immigration status, it is important to note that in order to have information in the Medical Birth Registry, an individual must be born in Denmark, and therefore, we could not include any first-generation immigrants, which is a limitation of the study. A previous study using register data from the Danish population (cohort of 2,486,646 million individuals born in Denmark from 1955 to 1993) showed an elevated relative risk of schizophrenia in offspring of immigrants compared to individuals with both parents born in Denmark [18]. Our sample differs, as we included all patients with a schizophrenia spectrum diagnosis and a control group matched on age, sex, and parental socioeconomic status. Moreover, our study covers a later time period than the above-mentioned study (1973–2018). The current findings could potentially be a result of the "healthy immigrant effect," a term used to describe the observation that immigrants' health is sometimes better than that of comparable native-born people at the time of migration [42]. This may be due to increasingly narrowing immigration laws in Denmark making it more difficult to migrate, and thus, only the more well-functioning and healthy part of a population manage to migrate. This would result in a positive selection bias, skewing the immigration population towards individuals who are less likely to develop a psychiatric illness. On a more speculative note, our finding may be explained by cultural differences making individuals from other cultural backgrounds less likely to seek psychiatric help in Denmark due to the risk of family shame or stigmatization, which would result in a larger number of unidentified cases.

As a measure of urbanicity, we included the population density of the birthplace. We did not have available information on where the individuals actually grew up or how much time they spent living in different areas, which is a limitation of the study. A Danish study showed that, in general, the more years spent living in areas with a higher degree of urbanisation, the higher the risk of schizophrenia [43]. Nevertheless, previous studies from Denmark have indicated that urbanicity of the birthplace itself is also linked to an increased risk of schizophrenia [9, 43, 44]. Compared to these studies, our cohort covers a later study period (1973–2018 compared to 1950–1998, 1956–1983, and 1955–2005, respectively). The contradictory findings may thus suggest a change in this variable as a risk factor in recent years. This could be explained by a shift in the Danish population during the last decades with fewer people now living in rural areas [45, 46]. In line with this, it is important to note that we did observe a highly skewed number of individuals in the three categories of urbanisation, with the majority of people falling into the medium category (10,000–100,000 inhabitants) and only very few in the lowest category (<10,000 inhabitants) (see Table 1). Nevertheless, we did observe the same pattern of higher population density being protective for schizophrenia risk when including continuous rather than categorical measures (Supplementary material Table S2), although the observed effect was small.

We only observed a significant interaction between parental history of psychiatric illness and immigration status, indicating that the protective effect of being a second-generation immigrant is higher in individuals with a parental history of psychiatric illness compared to individuals without. We observed no significant interactions between the remaining variables that conferred risk for schizophrenia, suggesting a cumulative rather than interactive effect in the presence of multiple of these factors.

https://doi.org/10.1192/j.eurpsy.2024.1774 Published online by Cambridge University Press

We observed no differences between males and females in terms of which of the included risk factors reached significance, suggesting that similar early risk factors are important for the development of schizophrenia across sexes. Nevertheless, for parental history of psychiatric illness, we observed a slightly higher odds ratio for males compared to females. In contrast, for low birth weight and high paternal age, we observed higher odds ratios for females. A stronger association between schizophrenia and high paternal age for females has also been shown in a previous Danish register study [11].

For the second hypothesis, we expected the early risk factors to be particularly important for an earlier age of onset resulting in a diagnosis before turning 18 years old. In this cohort, one out of every five patients was classified as an early-onset case. Here, we showed that female sex and parental history of psychiatric illness were the strongest predictors of receiving the diagnosis during childhood or adolescence compared to as an adult. This is in line with a previous study from our group demonstrating a stronger familial disposition for schizophrenia in females with an earlier illness onset [26]. Moreover, although the male sex generally has been shown to increase the risk of schizophrenia [14], a recent Danish register study showed a higher incidence for females in the youngest age group covering individuals up to 18 years old [47]. Higher birth weight, higher paternal age, and higher population density of birthplace also increased the risk of receiving the diagnosis early, although the observed effects were relatively small.

When restricting the child onset group to very early onset cases, we observed similar findings, although the effect of sex was no longer significant, possibly due to power issues. The significant effect of parental history of psychiatric illness showed higher OR for the very early onset cases, indicating a stronger familial load for this group consistent with previous findings [2].

Finally, for the third hypothesis, we expected a family history of psychiatric illness (as a proxy for the genetic effects) to be particularly important for schizophrenia risk and age of diagnosis. In line with this hypothesis, a family history of psychiatric illness showed the strongest impact in our models comparing patients with schizophrenia to HCs. Moreover, after female sex, a family history of psychiatric illness was the strongest predictor of receiving the diagnosis during childhood or adolescence. Nevertheless, family history of psychiatric illness may not be an appropriate proxy for the genetic effects per se [48], and it is difficult to disentangle the specific influences included in this variable. However, in our models, this variable was a robust predictor of both psychosis risk and age of diagnosis and could be easily included in routine assessments during pregnancy.

The current findings should be considered within the strengths and limitations of the study. The Danish Psychiatric Central Research Register is considered representative of patients with schizophrenia spectrum disorder, given that the number of privately treated patients is minimal [31]. Studies have demonstrated high validity for schizophrenia spectrum diagnoses [49, 50], including for child-onset cases [51]. The Medical Birth Registry was the primary focus of this study, and the basic information such as the early risk factors included here are considered of high validity. Some of these data are drawn directly from the Civil Registration System, which has very high validity, and others are mandatory to report when the newborn is registered after delivery [30]. The use of data collected at the time of birth limits the potential impact of recall bias when examining exposures occurring many years prior to the time of diagnosis. Moreover, the use of register data provides a mean to obtain information on rare disease outcomes such as schizophrenia in a large, highly representative sample, thus limiting the risk of recruitment bias. The large sample size allowed for adequate power to include multiple variables simultaneously as well as to apply more complex machine-learning techniques. By including multiple risk factors concurrently, we could account for intercorrelations and estimate the relative importance of each factor against each other.

On the other hand, the study has several limitations. The register data included here do not cover several important factors known to increase the risk of schizophrenia [6, 8], such as pregnancy and birth complications [52], premorbid intelligence [53], childhood trauma and cannabis use [7, 54]. Moreover, we estimated the genetic effect based on the family history of psychiatric illness rather than PRS [55]. In addition, due to the use of data from the medical birth registry, first-generation immigrants, as well as individuals who, for various reasons, may not access the healthcare service, are not included. Finally, while the use of register data facilitated the inclusion of a large and representative sample, for some of the variables, we did observe a high degree of missing data, which could confound the findings. However, using two different types of imputation did not change the overall results.

In conclusion, the extensive nationwide register data provided an opportunity to weigh the relative impact of multiple early factors on disease risk and age of diagnosis. Our findings indicate that the pathways to schizophrenia spectrum disorders are complex and most likely a culmination of multiple risk exposures acting in combination, each contributing with subtle effects [16, 56]. The current findings may be applied to guide the focus of intervention strategies to identify children at high risk, which from a health policy perspective, is crucial to prevent, delay or attenuate the impact of schizophrenia spectrum disorders. Routine assessments of the most influential risk factors, i.e., parental history of psychiatric illness, low birth weight and high paternal age, could readily be incorporated into clinical practise, facilitating individual risk stratification. Future studies should explore how these early risk factors influence prognostic outcomes such as cognitive impairments or treatment response as well as the more specific disturbances associated with schizophrenia, including different symptom domains.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1192/j.eurpsy.2024.1774.

Data availability statement. The register data included in this study is available from Statistics Denmark.

Acknowledgements. The authors would like to thank register consultant Marie Kruse for assistance with register data and Mikkel Erlang Sørensen for graphical assistance.

Financial support. This work was supported by a postdoctoral grant from the Mental Health Services in the Capital Region of Denmark (CKL).

Competing interest. BYG has been the leader of a Lundbeck Foundation Centre of Excellence for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS) (Jan 2009 to Dec 2021), which was partially financed by an independent grant from the Lundbeck Foundation based on the international review and partially financed by the Mental Health Services in the Capital Region of Denmark, the University of Copenhagen, and other foundations. All grants are the property of the Mental Health Services in the Capital Region of Denmark and administrated by them. BHE has received lecture fees and/or is part of Advisory Boards of Bristol-Myers Squibb, Eli Lilly, and Company, Janssen-Cilag, Otsuka Pharma Scandinavia AB, Takeda Pharmaceutical Company and Lundbeck Pharma A/S. The remaining authors report no conflicts of interest.

References

[1] Murray RM, Bhavsar V, Tripoli G, Howes O. 30 Years on: How the Neurodevelopmental Hypothesis of Schizophrenia Morphed into the Developmental Risk Factor Model of Psychosis. Schizophr Bull. 2017;43:1190–6. https://doi.org/10.1093/schbul/sbx121.

[2] De Berardis D, De Filippis S, Masi G, Vicari S, Zuddas A. A neurodevelopment approach for a transitional model of early onset schizophrenia. Brain Sci. 2021;11:1–16. https://doi.org/10.3390/brainsci11020275.

[3] Kahn RS, Sommer IE, Murray RM, Meyer-Lindenberg A, Weinberger DR, Cannon TD, et al. Schizophrenia. Nat Rev Dis Primers. 2015;1. https://doi.org/10.1038/nrdp.2015.67.

[4] Hilker R, Helenius D, Fagerlund B, Skytthe A, Christensen K, Werge TM, et al. Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. Biol Psychiatry. 2018;83:492–8. https://doi.org/10.1016/j.biopsych.2017.08.017.

[5] Cardno AG, Gottesman II. Twin studies of schizophrenia: From bow-and-arrow concordances to star wars Mx and functional genomics. Am J Med Gen. 2000;97:12–7. https://doi.org/10.1002/(SICI)1096-8628(200021)97:1<12::AID-AJMG3>3.0.CO;2-U.

[6] Radua J, Ramella-Cravaro V, Ioannidis JPA, Reichenberg A, Phiphopthatsanee N, Amir T, et al. What causes psychosis? An umbrella review of risk and protective factors. World Psychiatry. 2018;17:49–66. https://doi.org/10.1002/wps.20490.

[7] Lemvigh C, Brouwer R, Hilker R, Anhøj S, Baandrup L, Pantelis C, et al. The relative and interactive impact of multiple risk factors in schizophrenia spectrum disorders: A combined register-based and clinical twin study. Psychol Med. 2021. https://doi.org/10.1017/S0033291721002749.

[8] Matheson SL, Shepherd AM, Laurens KR, Carr VJ. A systematic meta-review grading the evidence for non-genetic risk factors and putative antecedents of schizophrenia. Schizophr Res. 2011;133:133–42. https://doi.org/10.1016/j.schres.2011.09.020.

[9] Pedersen CB, Mortensen PB. Family history, place and season of birth as risk factors for schizophrenia in Denmark: A replication and reanalysis. Br J Psychiatry. 2001;178:46–52. https://doi.org/10.1192/bjp.179.1.46.

[10] Larsen JK, Bendsen BB, Foldager L, Munk-Jørgensen P. Prematurity and low birth weight as risk factors for the development of affective disorder, especially depression and schizophrenia: A register study. Acta Neuropsychiatr. 2010;22:284–91. https://doi.org/10.1111/j.1601-5215.2010.00498.x.

[11] Byrne M, Agerbo E, Ewald H, Eaton WW, Mortensen PB. Parental Age and Risk of Schizophrenia: A Case-control Study. Arch Gen Psychiatry. 2003;60:673–8.

[12] Cantor-Graae E, Pedersen CB. Full spectrum of psychiatric disorders related to foreign migration: A danish population-based cohort study. JAMA Psychiatry. 2013;70:427–35. https://doi.org/10.1001/jamapsychiatry.2013.441.

[13] Nielsen SM, Toftdahl NG, Nordentoft M, Hjorthoj C. Association between alcohol, cannabis, and other illicit substance abuse and risk of developing schizophrenia: A nationwide population based register study. Psychol Med. 2017;47:1668–77. https://doi.org/10.1017/S0033291717000162.

[14] Thorup A, Waltoft BL, Pedersen CB, Mortensen PB, Nordentoft M. Young males have a higher risk of developing schizophrenia: A Danish register study. Psychol Med. 2007;37:479–84. https://doi.org/10.1017/S0033291707009944.

[15] Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "Just the Facts" What we know in 2008. 2. Epidemiology and etiology. Schizophr Res. 2008;102:1–18. https://doi.org/10.1016/j.schres.2008.04.011.

[16] Davis J, Eyre H, Jacka FN, Dodd S, Dean O, McEwen S, et al. A review of vulnerability and risks for schizophrenia: Beyond the two hit hypothesis. Neurosci Biobehav Rev. 2016;65:185–94. https://doi.org/10.1016/j.neubiorev.2016.03.017.

[17] Zwicker A, Denovan-Wright EM, Uher R. Gene-environment interplay in the etiology of psychosis. Psychol Med. 2018;48:1925–36. https://doi.org/10.1017/S003329171700383X.

[18] Sørensen HJ, Nielsen PR, Pedersen CB, Benros ME, Nordentoft M, Mortensen PB. Population impact of familial and environmental risk factors for schizophrenia: A nationwide study. Schizophr Res. 2014;153:214–9. https://doi.org/10.1016/j.schres.2014.01.008.

[19] Padmanabhan JL, Shah JL, Tandon N, Keshavan MS. The "polyenviromic risk score": Aggregating environmental risk factors predicts conversion to psychosis in familial high-risk subjects. Schizophr Res. 2017;181:17–22. https://doi.org/10.1016/j.schres.2016.10.014.

[20] Stepniak B, Papiol S, Hammer C, Ramin A, Everts S, Hennig L, et al. Accumulated environmental risk determining age at schizophrenia onset: A deep phenotyping-based study. Lancet Psychiatry. 2014;1:444–53. https://doi.org/10.1016/S2215-0366(14)70379-7.

[21] Tandon R, Nasrallah HA, Keshavan MS. Schizophrenia, "just the facts" 4. Clinical features and conceptualization. Schizophr Res. 2009;110:1–23. https://doi.org/10.1016/j.schres.2009.03.005.

[22] Pedersen CB, Mors O, Bertelsen A, LindumWaltoft B, Agerbo E, McGrath JJ, et al. A comprehensive nationwide study of the incidence rate and lifetime risk for treated mental disorders. JAMA Psychiatry. 2014;71:573–81. https://doi.org/10.1001/jamapsychiatry.2014.16.

[23] Hollis C. Adult outcomes of child- and adolescent-onset schizophrenia: Diagnostic stability and predictive validity. Am J Psychiatry. 2000;157:1652–9. https://doi.org/10.1176/appi.ajp.157.10.1652.

[24] Juola P, Miettunen J, Veijola J, Isohanni M, Jääskeläinen E. Predictors of short- and long-term clinical outcome in schizophrenic psychosis - the Northern Finland 1966 Birth Cohort study. European Psychiatry. 2013;28:263–8. https://doi.org/10.1016/j.eurpsy.2011.11.001.

[25] Chen L, Selvendra A, Stewart A, Castle D. Risk factors in early and late onset schizophrenia. Compr Psychiatry. 2018;80:155–62. https://doi.org/10.1016/j.comppsych.2017.09.009.

[26] Hilker R, Helenius D, Fagerlund B, Skytthe A, Christensen K, Werge TM, et al. Is an Early Age at Illness Onset in Schizophrenia Associated With Increased Genetic Susceptibility? Analysis of Data From the Nationwide Danish Twin Register. EBioMedicine. 2017;18:320–6. https://doi.org/10.1016/j.ebiom.2017.04.002.

[27] Schelin EM, Munk-Jørgensen P, Olesen A V., Gerlach J. Regional differences in schizophrenia incidence in Denmark. Acta Psychiatr Scand. 2000;101:293–9. https://doi.org/10.1034/j.1600-0447.2000.101004293.x.

[28] Byrne M, Agerbo E, Bennedsen B, Eaton WW, Mortensen PB. Obstetric conditions and risk of first admission with schizophrenia: A Danish national register based study. Schizophr Res. 2007;97:51–9. https://doi.org/10.1016/j.schres.2007.07.018.

[29] Thygesen LC, Daasnes C, Thaulow I, Brønnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: Structure, access, legislation, and archiving. Scand J Public Health. 2011;39:12–6. https://doi.org/10.1177/1403494811399956.

[30] Bliddal M, Broe A, Pottegård A, Olsen J, Langhoff-Roos J. The Danish medical birth register. Eur J Epidemiol. 2018;33:27–36. https://doi.org/10.1007/s10654-018-0356-1.

[31] Mors O, Perto GP, Mortensen PB. The Danish Psychiatric Central Research Register. Scand J Public Health. 2011;39:54–7. https://doi.org/10.1177/1403494810395825.

[32] Schmidt M, Alba S, Schmidt J, Sandegaard JL, Ehrenstein V, Pedersen L, et al. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol. 2015;7:449–90. https://doi.org/10.2147/CLEP.S91125.

[33] Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. Eur J Epidemiol. 2014;29:541–9. https://doi.org/10.1007/s10654-014-9930-3.

[34] Davies G, Welham J, Chant D, Torrey EF, McGrath J. A systematic review and meta-analysis of northern hemisphere season of birth studies in schizophrenia. Schizophr Bull. 2003;29:587–93. https://doi.org/10.1093/oxfordjournals.schbul.a007030.

[35] Torrey EF, Buka S, Cannon TD, Goldstein JM, Seidman LJ, Liu T, et al. Paternal age as a risk factor for schizophrenia: How important is it? Schizophr Res. 2009;114:1–5. https://doi.org/10.1016/j.schres.2009.06.017.

[36] Bourque F, Van Der Ven E, Malla A. A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. Psychol Med. 2011;41:897–910. https://doi.org/10.1017/S0033291710001406.

[37] The American College of Obstetricians and Gynecologists. Committee opinion: the Apgar score. Obstet Gynecol. 2015;126:691–2.

https://doi.org/10.1192/j.eurpsy.2024.1774 Published online by Cambridge University Press

[38] Kuo CL, Duan Y, Grady J. Unconditional or conditional logistic regression model for age-matched case–control data? Front Public Health 2018;6. https://doi.org/10.3389/fpubh.2018.00057.
[39] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res. 2011;12:2825–30.
[40] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 13–17 Aug, 2016, p. 785–94. https://doi.org/10.1145/2939672.2939785.
[41] Vassos E, Pedersen CB, Murray RM, Collier DA, Lewis CM. Meta-analysis of the association of urbanicity with schizophrenia. Schizophr Bull. 2012;38:1118–23. https://doi.org/10.1093/schbul/sbs096.
[42] McDonald JT, Kennedy S. Insights into the "healthy immigrant effect": Health status and health service use of immigrants to Canada. Soc Sci Med. 2004;59:1613–27. https://doi.org/10.1016/j.socscimed.2004.02.004.
[43] Pedersen CB, Mortensen PB. Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. Arch Gen Psychiatry. 2001;58.
[44] Laursen TM, Munk-Olsen T, Nordentoft M, Mortensen PB. A comparison of selected risk factors for unipolar depressive disorder, bipolar affective disorder, schizoaffective disorder, and schizophrenia from a Danish population-based cohort. J Clin Psychiatry. 2007;68.
[45] Statistics Denmark. Byopgørelsen 1. januar 2023: Færre bor på landet. 2023.
[46] Statistics Denmark. Byopgørelsen 1. januar 2015 Færre på landet-flere i de større byer. 2015.
[47] Kühl JOG, Laursen TM, Thorup A, Nordentoft M. The incidence of schizophrenia and schizophrenia spectrum disorders in Denmark in the period 2000–2012. A register-based study. Schizophr Res. 2016;176:533–9. https://doi.org/10.1016/j.schres.2016.06.023.
[48] Mars N, Lindbohm J V., della Briotta Parolo P, Widén E, Kaprio J, Palotie A, et al. Systematic comparison of family history and polygenic risk across 24 common diseases. Am J Hum Genet. 2022;109:2152–62. https://doi.org/10.1016/j.ajhg.2022.10.009.
[49] Uggerby P, Østergaard SD, Røge R, Correll CU, Nielsen J. The validity of the schizophrenia diagnosis in the Danish Psychiatric Central Research Register is good. Dan Med J. 2013;60.
[50] Löffler W, Häfner H, Fätkenheuer B, Maurer K, Riecher-Rössler A, Lützhøft J, et al. Validation of Danish case register diagnosis for schizophrenia. Acta Psychiatr Scand. 1994;90:196–203. https://doi.org/10.1111/j.1600-0447.1994.tb01577.x.
[51] Vernal DL, Stenstrøm AD, Staal N, Christensen AMR, Ebbesen C, Pagsberg AK, et al. Validation study of the early onset schizophrenia diagnosis in the Danish Psychiatric Central Research Register. Eur Child Adolesc Psychiatry. 2018;27:965–75. https://doi.org/10.1007/s00787-017-1102-z.
[52] Cannon M, Jones PB, Murray RM. Obstetric Complications and Schizophrenia: Historical and Meta-Analytic Review. Am J Psychiatry. 2002;159:1080–92. https://doi.org/10.1176/appi.ajp.159.7.1080.
[53] Khandaker GM, Barnett JH, White IR, Jones PB. A quantitative meta-analysis of population-based studies of premorbid intelligence and schizophrenia. Schizophr Res. 2011;132:220–7. https://doi.org/10.1016/j.schres.2011.06.017.
[54] Harley M, Kelleher I, Clarke M, Lynch F, Arseneault L, Connor D, et al. Cannabis use and childhood trauma interact additively to increase the risk of psychotic symptoms in adolescence. Psychol Med. 2010;40:1627–34. https://doi.org/10.1017/S0033291709991966.
[55] Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421–7. https://doi.org/10.1038/nature13595.
[56] Stilo SA, Murray RM. Non-genetic factors in schizophrenia. Curr Psychiatry Rep. 2019;21(10):100. https://doi.org/10.1007/s11920-019-1091-3. PMID: 31522306; PMCID: PMC6745031.

https://doi.org/10.1192/j.eurpsy.2024.1774 Published online by Cambridge University Press