**Supplementary Material**

**Literature Review**

It has been shown in the Western world, particularly in Europe and the United States, that ethnic minority groups are at greater risk of developing psychosis and related disorders such as schizophrenia and other severe mental illnesses. Factors such as substance use, misdiagnosis, cultural bias and socioeconomic status have not been shown to adequately explain such elevated risk. This has lead researchers to explore social determinants in the development of psychosis, including identity, discrimination, and disadvantage as potential explanatory mechanisms. Relating to these social causes, it has been found that living in areas with a larger proportion of one’s own ethnic group can mitigate the increased risk of psychosis in ethnic minority groups, described as the ethnic density effect. However, this effect may not be consistent across different ethnic minority groups. In particular, survey studies have shown an absence of protective ethnic density effects for Pakistani groups in the UK, which is not seen in other South Asian groups.

We searched PubMed in April 2023 using the following search strategy to identify evidence exploring ethnic density and psychosis outcomes, with no restriction on publication date: ((((((("psychosis"[Title]) OR ("schizophrenia"[Title])) OR ("psychotic disorder"[Title])) OR ("severe mental illness"[Title])) AND ("ethnic density"[Title]))) OR ("ethnic density effect"[Title]). This search identified 18 results, including two published systematic reviews with meta-analyses. The reviews indicated a consistent absence of protective effects of ethnic density for the Pakistani population in the UK, as well as heterogeneity of effects across different ethnic minority groups. It was also highlighted that findings regarding ethnic density effects vary depending on study methodology, particularly choice of outcome measure, and the size of geographical area used to compare area level effects. It has been demonstrated that using clinically relevant measures of psychosis outcome, and also smaller geographical units (such as Lower Super Output Areas in the UK census), can show ethnic density effects which may otherwise be obscured by self-report measures and larger geographical units such as electoral wards. It is unclear whether the absence of a protective ethnic density effect in the UK Pakistani population relates to such methodological concerns.

**Supplementary Analyses**

**Analysis of Ethnic Density and FEP in White Sample**

|  |  |  |  |
| --- | --- | --- | --- |
| **Table S1. Model assessing factors influencing incidence rates of FEP in White patients (n=364)** | | | |
| **Variable** | | **Adjusted IRRa** | **95% CI** |
| Gender (female=0/male=1) | | 1·14\* | 1·02-1·27 |
| Age | | 1·02 | 0·93-1·13 |
| Area Deprivation (IMD) | | 1·01\* | 1·00-1·01 |
| Ethnic Density Quartile | | | |
|  | 4th Quartile (98.59%) | - | - |
|  | 3rd Quartile (97.13%) | 1·02 | 0·85-1·24 |
|  | 2nd Quartile (93.32%) | 0·86 | 0·74-1·00 |
|  | 1st Quartile (least dense: 59.86%) | 1·24\* | 1·01-1·54 |
| aAdjusted for age, gender and area level deprivation  \**p<·05* | | | |

**FEP and ARMS Combined Sample Analyses**

For the analysis, a total of 575 patients (473 White, 102 Pakistani) were included. Of these, 209 were female and 366 were male, with a mean age of 25 (range: 13-63) years. The 575 cases of FEP and ARMS combined were identified over a total of 751,266 person-years follow up. A crude incidence rate of 71 per 100,000 person-years was found for the White category, and a rate of 115 per 100,000 for the Pakistani category, with an unadjusted Incidence Rate Ratio IRR of 1·60 (95% CI 1·20-2·15, p = ·002). Adjusting for age and gender gave an IRR of 1·62 (95% CI 1·23-2·15 p = ·001). The final adjustment for area level deprivation gave an IRR of 1·52 (95% CI 1·13-2·04, p = ·006), showing significantly higher rates of FEP and ARMS for the Pakistani group relative to the White population in the East Lancashire population.

|  |  |  |
| --- | --- | --- |
| **Table S2. Incidence of FEP for Pakistani patients compared to White patients in areas of Low and High Ethnic Density** | | |
| **Ethnic Density** | **Adjusted IRR** | **95% CI** |
| Low (0%-11·79%) | 13·71\* | 9·03-20·80 |
| High (11·80%-73%) | 0·86 | 0·63-1·18 |
| aAdjusted for age, gender and area level deprivation  *\*p<·001* |  |  |

Once Pakistani ethnic density was included in the model, a significant interaction between ethnicity and ethnic density was found (p < ·001). IRRs were compared between areas with high ethnic density and low ethnic density (see Table S2), while adjusting for age, gender, and deprivation, to examine the interaction. In the areas with high ethnic density, incidence rates of FEP and ARMS between the Pakistani and White categories did not significantly differ (IRR 0·86, 95% CI 0·63-1·18, p = ·358). Conversely, in areas of low ethnic density, the Pakistani category had significantly higher rates of compared to that of the White category (IR 13·71, 95% CI 9·03-20·80, p < ·001).

|  |  |  |  |
| --- | --- | --- | --- |
| **Table S3. Model assessing factors associated with incidence rates of FEP+ARMS in Pakistani patients (n=102)** | | | |
| **Variable** | | **Adjusted IRRa** | **95% CI** |
| Gender (female=0/male=1) | | 1·01 | 0·78-1·30 |
| Age | | 1·57\* | 1·22-2·02 |
| Area Deprivation (IMD) | | 1·00 | 0·99-1·01 |
| Ethnic Density Quartile | | | |
|  | 4th Quartile (most dense: 60·44%) | - | - |
|  | 3rd Quartile (38·51%) | 1·61\* | 1·12-2·32 |
|  | 2nd Quartile (19·98%) | 2·74\*\* | 2·10-3·58 |
|  | 1st Quartile (least dense: 2·54%) | 11·66\*\* | 7·82-17·97 |
| aAdjusted for age, gender and area level deprivation  \**p<·05, \*\*p<·001* | | | |

A further model assessing Pakistani patients only (see Table S3), with equally sized quartiles from high ethnic density to low ethnic density, was used to examine the ethnic density effect in more detail. This model shows that those in the lowest density quartile have a markedly higher risk of FEP and ARMS when compared to those in the highest density quartile (IRR 11·66, 95% CI 7·82-17·97, p < ·001). The Pakistani group also showed higher rates in the second (IRR 2·74, 95% CI 2·10-3·58, p < ·001) and third (IRR 1·61, 95% CI 1·12-2·32, p = ·011) lowest density quartiles compared to the highest density. In contrast, the equivalent model for White patients did not show evidence of an ethnic density effect (see Table S4).

|  |  |  |  |
| --- | --- | --- | --- |
| **Table S4. Model assessing factors influencing incidence rates of FEP+ARMS in White patients (n=473)** | | | |
| Variable | | Adjusted IRRa | 95% CI |
| Gender (female=0/male=1) | | 1·18\* | 1·06-1·31 |
| Age | | 1.00 | 0·92-1·09 |
| Area Deprivation (IMD) | | 1·01\*\* | 1·00-1·01 |
| Ethnic Density Quintile | | | |
|  | 4th Quartile (98.61%) | - | - |
|  | 3rd Quartile (97.15%) | 0·92 | 0·78-1·09 |
|  | 2nd Quartile (93.79%) | 0·85\* | 0·73-0·98 |
|  | 1st Quartile (least dense: 59.60%) | 1·00 | 0·84-1·19 |
| aAdjusted for age, gender and area level deprivation  \**p<·05 \*\*p<·001* | | | |

**Table S5: STROBE Statement**

|  |  |  |
| --- | --- | --- |
|  | Item No | Recommendation |
| **Title and abstract** | 1 | (*a*) Indicate the study’s design with a commonly used term in the title or the abstract |
| (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found |
| Introduction | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses |
| Methods | | |
| Study design | 4 | Present key elements of study design early in the paper |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection |
| Participants | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants |
| (*b*)*Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study*—For matched studies, give matching criteria and the number of controls per case |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable |
| Data sources/ measurement | 8\* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group |
| Bias | 9 | Describe any efforts to address potential sources of bias |
| Study size | 10 | Explain how the study size was arrived at |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding |
| (*b*) Describe any methods used to examine subgroups and interactions |
| (*c*) Explain how missing data were addressed |
| (*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed  *Case-control study*—If applicable, explain how matching of cases and controls was addressed  *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy |
| (*e*) Describe any sensitivity analyses |

Continued on next page

|  |  |  |
| --- | --- | --- |
| Results | | |
| Participants | 13\* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed |
| (b) Give reasons for non-participation at each stage |
| (c) Consider use of a flow diagram |
| Descriptive data | 14\* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders |
| (b) Indicate number of participants with missing data for each variable of interest |
| (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) |
| Outcome data | 15\* | *Cohort study*—Report numbers of outcome events or summary measures over time |
| *Case-control study—*Report numbers in each exposure category, or summary measures of exposure |
| *Cross-sectional study—*Report numbers of outcome events or summary measures |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included |
| (*b*) Report category boundaries when continuous variables were categorized |
| (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses |
| Discussion | | |
| Key results | 18 | Summarise key results with reference to study objectives |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results |
| Other information | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based |

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.