Synopsis no.: S2.55

Preliminary title: Re-evaluating setting-level variance in the incidence of psychotic disorders in the EU-GEI WP2 study

Contact info for the person(s) proposing the synopsis

Name: **James B Kirkbride**

Partner no: **UCL**

e-mail address: **j.kirkbride@ucl.ac.uk**

Publication category: ???

Working and writing group: Kirkbride, Andleeb, Jongsma, Morgan, Di Forti, Murray + any other WP2 member who would like to actively contribute to the synopsis

Work Packages involved: WP2

Partners involved from whom candidate co-authors (*additional to working and writing group*) should be nominated:

Open to any WP2 member who would like to actively contribute to the study

Objectives (scientific background, hypothesis, methods, and expected results):

1. Scientific Background/ objectives

The EU-GEI WP2 authorship group has previously published two major papers investigating reasons for variation in the incidence of psychotic disorders by setting in the study. In the first paper, Jongsma et al[1], reported the overall incidence of psychotic disorders, including non-affective and affective psychotic disorders separately, by setting and country. Using multilevel Poisson regression, they investigated whether crude and age-sex-ethnicity standardised incidence rates varied between all 17 settings according to setting latitude, % owner occupancy, % unemployment, % living alone and population density. In mutually adjusted models, they only found an association between a lower proportion of the population who owned their own home (a proxy marker for socioeconomic deprivation) and higher incidence rate of psychotic disorders.

In a second investigation, Di Forti and colleagues[2] investigated the contributing role of cannabis use to risk of psychotic disorder, using both the case-control and incidence data available in WP2 of the EU-GEI study. In their incidence analyses, Di Forti et al inspected the correlation between age-sex-ethnicity adjusted incidence rates and the proportion of high potency and daily cannabis use amongst controls in 11 of 17 settings in the study, reporting correlations of between 0.7-0.8 ($p<0.02$).

These two studies suggest possibly different, independent or joint mechanisms may contribute to variance in the incidence of psychotic disorders between settings. However, Jonsgma et al[1] did not include the markers

of cannabis used by Di Forti et al[2] in their original analyses, and Di Forti et al[2] did not include the potential confounding role of the social markers included by Jongsma et al. [1] One further issue to address is missing data from the original cannabis experiences questionnaire in the case-control study, which led to the omission of some settings in the analyses by Di Forti et al. [2] It may be possible to use multiple imputation to recover some of this missing information to understand the unbiased association between population-level cannabis use and the incidence of psychotic disorders across different settings. For these reasons, this issue warrants further investigation and re-analysis.

2. Hypotheses

We hypothesise that:

i. There will be substantive between-setting variance in the incidence of all psychotic disorders and non-affective psychotic disorders in null and individual-level adjusted models (for age-sex-ethnicity). This variation will be less pronounced for the affective psychoses (consistent with the literature) and for all outcomes, will be attenuated after full modelling of setting level covariates

ii. Socioeconomic deprivation – indexed via % owner occupancy – will contribute independently to variance in the incidence of psychotic disorders than population-level daily cannabis use, and more substantially than high potency cannabis use at the area-level.

3. Methods

Data sources
Incidence data will be provided from 16 sites of WP2 in the EU-GEI study, as per the methods detailed in Jongsma et al. Data from the central Paris site will be excluded as no corresponding controls were collected at this site to provide estimates of the prevalence of cannabis use. Case control data on cannabis use from the cannabis experiences questionnaire and auxiliary case-control data to inform multiple imputation methods will be obtained from case-control data in WP2.

Outcomes
We will study non-affective and affective psychotic disorders separately, and as a joint outcome, using the classification system previously defined and reported in WP2.

Exposures
Setting-level % owner-occupancy, % single person households, % unemployed, latitude, population density and the proportion of controls who reported daily cannabis use, or using high potency cannabis will be considered as our main exposure variables.

Confounders
Individual level age group, sex and ethnicity (as per Jongsma et al) will be considered as covariates.

Denominator data
Denominator data from each setting will be provided from the relevant Censuses in each country, nearest the time of case ascertainment as detailed in Jongsma et al.

Statistical analyses
First, we will inspect the proportion of missing data by setting on responses to the CEQ needed to generate our daily use and high potency cannabis variables. Second, we will then use multiple imputation via chained equations [MICE] to predict missing cannabis items required to generate our daily use and high potency cannabis variables in controls. We will include suitable auxiliary variables in our model, including setting, country, ethnicity/migrant status, age, sex, marital status, family history of mental illness, education level, socioeconomic status and any other relevant variables. We will generate two sets of cannabis use variables at the setting level: (i) the non-imputed (i.e. complete case) proportions by setting for daily and high potency use, and (ii) the imputed proportions of the same variables. Third, we will investigate correlations between setting-level variables and the crude and age-sex-ethnicity standardised rates of each outcome across settings.

Fourth, we will conduct univariable and multivariable Poisson analyses to test our main hypotheses. Each setting-level exposure will be entered into the multivariable model in order of their univariable association with the incidence of a given outcome, assessed via AIC. Forward-fitting model building will be assessed via likelihood ratio test [LRT] until derivation of a final model. Multiply-imputed cannabis measures will be considered our primary measures on this variable, and compared to equivalent univariable and final multivariable models using the complete case variables.

References

1 Jongsma HE, Gayer-Anderson C, Lasalvia A, *et al.* Treated Incidence of Psychotic Disorders in the Multinational EU-GEI Study. *JAMA Psychiatry* 2018; **75**: 36–46.

2 Di Forti M, Quattrone D, Freeman TP, *et al.* The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-GEI): a multicentre case-control study. *The Lancet Psychiatry* 2019; **6**: 427–36.

Data needed for the study:

As described above. Raw and any available derived variables used in Jongsma et al or Di Forti et al will be used.

Plan for statistical analysis (overall strategy):

See above

Other analyses/methods:

N/A

Involvement of external Parties (non EU-GEI):

Humma Andleeb, UCL first-year PhD student (Supervisor: James Kirkbride)

IPR check:

????

Timeframe:

Immediate to end of 2021, plus any additional time needed for peer-reviewed publication.

Additional comments:

None.