Project Number: CC024

Title: Triangulating evidence from multiple control groups to study the causes and consequences of being born with cleft lip/palate

PI: Connor Moore

Affiliation: in Affiliation with the University of Bristol

Approval Date: April 2020

Scientific Outline:

The Cleft Collective Cohort Study has over 3000 children born with Cleft Lip and/or Palate (CLP) and was created to establish the causes and consequences of CLP. For this to be determined, the Cleft Collective cohort needs to be compared to a control data set. Presently, there are several birth cohorts to provide a resource for a control. However, using one data set alone leads to potential bias in any future comparisons. As a result, false differences could be attributed to the CLP population, which are instead due to a unique cohort-related factor in the control group.

The main aim of this research project is to triangulate evidence from several control sets, such as the ALSPAC, Born in Bradford, and the millennium cohort study. This triangulation is necessary to mitigate for any cohort specific biases that are in any one control group. This will be conducted by first finding differences such as age, year of birth or geographical differences between the Cleft Cohort and Control group. These variables will then be assessed for their direction and degree of bias on effect estimates in future comparisons. If it is found that these variables do not have a significant biasing effect, then evidence will be collated from the selected control groups and used for future comparisons to the Cleft Collective Cohort. This, therefore, will provide more confidence in any future findings showing differences in the CLP cohort from the control set are due to having CLP alone rather than another confounding variable. To pilot this study, we will be doing a singular comparison between the Millennium Cohort Study and the Cleft Collective Cohort. For this to be possible we need access to the cohort data which outlines the pre-natal and socioeconomic differences and/or similarities between the two cohorts. This would be possible to access with permission to extract data from the Baseline Questionnaires of the Birth Cohort and 5y Cohort. The variables will be summarised by calculating proportions (for categorical/binary variables) and means/medians and distributions (range, standard deviation, interquartile range) for continuous variables. Linear and logistic regression analyses will then be used to assess associations between prenatal exposures to offspring outcomes and cleft status using the cleft collective and each of the control groups sequentially. The effect estimates and confidence intervals will then be plotted to assess agreement between the different cohorts.