International Journal of Population Data Science

Journal Website: www.ijpds.org

A National Population-Based E-cohort of People with Psychosis (PsyCymru) Linkage of Phenotypical and Genetic Data to Routinely Collected Records

Lloyd, K[1], John, A[2], Craddock, N[3], Walters, J[4], Linden, D[5], Jones, I[4], Bentall, Department of Psychology University of Sheffield[6], Lyons, R[7], Lee, SC[1], and Owen, M[4,5,8]

[1]Swansea University Medical School

[2]Swansea University

[3]Department of Psychological Medicine and Neurology, Cardiff University

[4]National Centre for Mental Health, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University

[5]Division of Psychological Medicine and Clinical Neurosciences, Cardiff University

[6]Department of Psychology University of Sheffield

[7]Farr Institute, Swansea University Medical School David Ford, Swansea University

[8]Neuroscience and Mental Health Research Institute, Cardiff University

Introduction

PsyCymru was established to investigate the feasibility of linking a prospectively ascertained, well characterised (linked clinical cohort) of people with psychosis in Wales, UK with large amounts of anonymised routinely collected health record data. We are now additionally linking genetic data.

Objectives and Approach

PsyCymru aimed to create a research platform for psychosis research in Wales by establishing two cohorts. The first was a well-characterised clinically assessed cohort with genetic data. Consented individuals underwent structured interviews using well-validated questionnaires and gave blood sample for DNA extraction, sequencing, and candidate gene identification. This data was then linked to routinely collected health and social datasets with identity encryption. The second is a larger e-cohort of prevalent psychosis cases created using a validated algorithm applied to anonymised routine data. Both cohorts were tracked prospectively and retrospectively in the Secure Anonymised Information Linkage (SAIL) databank.

Results

In total, data from 958 individuals for the clinical cohort were imported to SAIL. Among these individuals, genetic data for 740 were analysed. The genetic data included robust loci for schizophrenia, pathogenic copy-number variations (CNVs) for various conditions (e.g., autism, intellectual disability, congenital malformations), polygenic risks scores for schizophrenia, as well as pathogenic/non-pathogenic duplications or deletions of chromosome spanning more than 500kb or 1Mb. For the e-cohort, 29,797 individuals were found having a psychosis diagnosis from primary and secondary care between 2004 to 2013. Social demographic data for both cohorts were also analysed based on sex, age, area deprivation, urbanicity, and employment status.

Conclusion/Implications

This unique platform pooled data together from multiple sources; linking clinical, psychological, biological, genetic, and health care factors to address assorted research questions. This resource will continue to expand over the coming years in size, breadth and depth of data, with continued recruitment and additional measures planned.

http://dx.doi.org/10.23889/ijpds.v3i4.706