Scottish Government Health Directorates Chief Scientist Office

CHIEF SCIENTIST OFFICE

FOCUS ON RESEARCH

Utilising Routinely Collected Electronic Medical Records to Predict Dementia

Researchers

Dr Wilfred Bonney, Dr Emily Jefferson, & Dr Alexander Doney

Aim

Dementia has catastrophic implications for affected individuals, their family and wider society. However, patients with dementia are challenging to study and identify for treatment because the diagnosis is sporadically and imprecisely recorded in electronic medical records (EMRs). This study aimed to develop and validate a phenotype algorithm (a set of rules based on using clinical and other patient characteristics) to identify patients who have developed dementia since being recruited into the GoDARTS bioresource (a large anonymous database of Scottish patients that includes genetic data).

Project Outline/Methodology

The study was carried out by:

- Exploiting key datasets within the GoDARTS bioresource to identify individuals who have had dementia.
- Developing an EMR-derived analysable dementia phenotype algorithm.
- Applying the algorithm iteratively on five selected datasets from the GoDARTS bioresource.
- Validating and optimising the algorithm with established genetic markers of disease risk such as the apolipoprotein E ε4 (ApoE4) single nucleotide polymorphism (SNP) and a weighted Genetic Risk Score (wGRS), calculated using 20 known susceptibility SNPs associated with Alzheimer's disease.
- Performing analyses to estimate the risk of dementia association with ApoE4 and wGRS.
- Measuring the performance of the algorithm.

Key Results

The algorithm identified 855 probable and 840 possible cases of dementia from the GoDARTS bioresoure of 18,190. Whereas the *Probable Cases* were defined as individuals with specific clinical codes for dementia, the *Possible Cases* were defined as individuals uniquely referred to "Old Age Psychiatry" clinic without any specific clinical codes for dementia. The overall incidence rate of dementia in the probable cases was 7.28 per 1000 person-years. The algorithm further classified the probable cases into six subtypes of dementia: Alzheimer's disease, Vascular dementia, Mixed dementia, Unspecified dementia, Other dementias, and Prescribing-based dementia. The algorithm yielded a positive predictive value (PPV) of 93.8%.

Conclusions

The study achieved its aim by demonstrating that it is possible to identify patients with dementia from large population-based bioresources using genetic instruments for validation and optimisation. The approach generated a dementia specific sub-cohort for the international research community to support translational research.

What does this study add to the field?

The study adds to the growing body of knowledge in Health Informatics literature that stricter or greater adherence to standardisation of routinely collected primary care data subsequently facilitate EMR-derived phenotypes in medical research.

Implications for Practice or Policy

A fundamental ramification of this study is to facilitate a dementia diagnosis algorithm to feed into primary care to help in identifying a population at increased risk who may need closer monitoring, as well as provide a method of enriching recruitment into key longitudinal studies in the field of Alzheimer's disease.

Where to next?

The EMR-derived dementia phenotype algorithm will be used as the foundation to build libraries of analysable disease-based phenotypes (e.g. Stroke, Myocardial infarction, and Congestive heart failure) for medical record linkage within the GoDARTS bioresource.

Further details from:

Dr Alexander Doney, NHS Consultant Physician (a.doney@dundee.ac.uk)