The Challenge:
Identifying patients with a particular condition from electronic health records (EHRs) offers the great potential to improve clinical trial recruitments, survival analysis, outcome predictions and many other medical studies. But the question is – how to do this?

Currently identifying patients with a particular condition can be extremely time-consuming and costly, while the criteria which characterise conditions, is buried within complex terminology across many areas in the records of a patient.

The Research:
To address these challenging issues, data analysts at the Farr Institute CIPHER have developed new techniques to automatically identify a group of patients with a particular condition from doctor’s electronic records (primary care). This study focused on the identification of rheumatoid arthritis (RA) as an example.

The study looked at 20,667 patients, linking via the SAIL databank (SAIL), doctor’s electronic health records to a secondary care (hospital) rheumatology records. SAIL is a world-class system that brings data together in a secure, trusted and confidential way.

The findings from the specialist rheumatologist consultants were used as the gold standard for developing a new technique, so that the influential codes of a medical condition identified in primary care records can accurately predict a diagnosis of the condition in secondary care health records.

The Results:
This study identified a final 8 read codes in the primary care (doctors) records, which become the potential predictors for identifying patients with rheumatoid arthritis from those without.

The Impact:
The data-driven method developed by the team provides a cost-effective and time saving tool for the NHS to reliably identifying patients with rheumatoid arthritis and helping clinicians in their decision-making process.

For further information visit: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0154515

Enquiries to Sarah Toomey, Communications Officer, Farr Institute CIPHER, s.toomey@swansea.ac.uk

Shang-Ming Zhou, Fabiola Fernandez-Gutierrez, Jonathan Kennedy, Roxanne Cooksey, Mark Atkinson and Sinead Brophy, Swansea University
Stefan Siebert, University of Glasgow; William Dixon and Terence O'Neill, University of Manchester
Ernest Choy, Cardiff University; Cathie Sudlow, University of Edinburgh
Spiros Denaxas, UCL London