THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

A tool to improve the efficiency and reproducibility of research using electronic health record databases

Citation for published version:
Al Sallakh, M, Rodgers, S, Lyons, R, Sheikh, A & Davies, G 2018, 'A tool to improve the efficiency and reproducibility of research using electronic health record databases', International Journal of Population Data Science, vol. 3, no. 2, pp. 71.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
International Journal of Population Data Science

OPEN ACCESS

International Journal of Population Data Science

A tool to improve the efficiency and reproducibility of research using electronic health record databases

Al Sallakh, M[1][*], Rodgers, S[1], Lyons, R[1], Sheikh, A[2], and Davies, G[1]

[1]Swansea University Medical School

[1]Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh

Background

Interrogation of electronic health record databases often involves time-consuming, manual, repetitive work in developing database queries. We developed a tool to automate this process.

Methods

We identified elementary approaches to query primary care data from the Secure Anonymised Information Linkage databank of Wales. We designed a web-based query builder that allows using combinations of these approaches as 'building blocks' to query complex variables. We created an R programme to automatically generate and execute the corresponding Structured Query Language queries.

Results

The tool allows data extraction using combinations of the following methods: event count (e.g., asthma prescriptions); code/date of earliest/latest event; code/date/value of the event of maximum/minimum value; and frequency of temporally constrained events. Query intervals could be fixed, dynamic, or individualised. The tool integrates with a codeset repository. Data extraction procedures and codesets are saved on a web server as versioned, shareable, and citable objects.

Conclusion

This versatile tool allows rapid and complex data extraction with minimal to no programming skills, reduces human errors, and improves research transparency and reproducibility.

Funding/Support

*Corresponding Author:
Email Address: m.a.alsallakh@swansea.ac.uk (M Al Sallakh)