Follow us @BoneJointRes

Freely available online OPEN ACCESS

EDITORIAL

Future registry research

DATABASE LINKAGE IN TRAUMA AND ORTHOPAEDICS

L. Farrow,
J. Evans

From University of Aberdeen, Aberdeen, UK

Cite this article: *Bone Joint Res* 2023;12(4):256–258.

Keywords: Data, Registry, Linkage

Trauma and Orthopaedics has long been established as a leader in the collection and analysis of high-quality patient data for the purposes of quality improvement and research. The scale and scope of registry data across several patient groups and procedures are the envy of colleagues around the world. The ability to link these data to other routinely held clinical information (e.g. patient health records held within the Hospital Episode Statistics (HES) data warehouse) allows us to enrich these data with important information that is not present in the original database, and would otherwise be a substantial burden to collect for every patient.

The main strength of mandatory national registries is the data completeness (or coverage) of the study population. The drawback, however, is that regular changing of the collected data (or minimum dataset) makes any analysis cumbersome and becomes a burden to those submitting data. As a result, if variables are not collected in the registry, we must look to other sources to supplement it when needed for a study. An example of this is if a researcher wants to investigate mortality outcomes of patients undergoing joint arthroplasties (included within the UK National Joint Registry (NJR)), while adjusting for specific comorbidities (not included within the NJR).

Requests to link datasets to routinely collected information (which is typically held by a national central organization, e.g. NHS Digital) are typically done through a formal application process. For instance, for the UK this is through the Data Access Request Service (DARS) in England and Wales, the Electronic Data Research and Innovation Service (eDRIS) in Scotland, and the Northern Ireland Statistics and Research Agency (NISRA) in Northern Ireland.

Specific registries may have pre-existing established linkage to routinely collected data (e.g. the NJR and HES), which means that applications can be made direct to the registry without having to go through the centralized service. Any application to link registry data to routinely collected healthcare data would typically be undertaken in combination with members of the registry committee or research team (such as the UK NJR research sub-committee), who will have specific expertise and understanding as to the vagaries of their particular dataset that are important for the linkage process.

As an example, there must be a common feature of both datasets (such as NHS number) that allows us to ensure that the data in each source relate to the same patient; people experienced in data linkage of individual datasets will understand how this is best achieved.

When datasets are linked, there are regulations that need to be followed: in particular, for any application to link data there are some key components of the General Data Protection Regulation (GDPR) and Data Protection Act 2018 (DPA) (which govern appropriate use of patient data) to be aware of.

Firstly, there is the role of the Data Controller (the organization(s)/person(s) with ultimate control of, and responsibility for, the data in question) and the Data Processor (the organization(s)/person(s) performing data processing (e.g. obtaining, recording, or holding data) on behalf of the Data Controller). Both are integral to any data application and will need to be clearly defined as part of the application process.

Secondly, there is the lawful basis for data processing. This is covered by Article 6 (1) of the GDPR, which sets out six potential reasons that can be used for lawful data processing.[1]

Correspondence should be sent to Luke Farrow; email: luke.farrow@abdn.ac.uk

doi: 10.1302/2046-3758.124.BJR-2023-0072

Bone Joint Res 2023;12(4):256–258.

Most healthcare-related linkage applications from NHS or academic institutions will be covered under Article 6 (1) –I - "performance of a task carried out in the public interest or in the exercise of official authority vested in the controller." Healthcare applications are also considered "special category data", which require further additional justification under Article 9 of the GDPR. In most cases this will be covered under Article 9 (2)(j) – where "processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes," or Article 9 (2) (h) – "processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3."

Another consideration is the common law duty of confidentiality. This governs the requirements for appropriate informed consent for the use of potentially identifiable confidential information. Given that most datasets use "pseudoanonymised" (the use of fictional identifiers to categorize individuals and link data together) data, the common law duty of confidentiality is applicable. In the case of truly anonymized data, consent is not required. The legal obligations specified in the common law duty of confidentiality (requirement for informed consent) for national datasets can be set aside through application to the Confidentiality Advisory Group (CAG), to allow application of Section 251 of the NHS Act 2006. A similar process also exists in Scotland through application to the Public Benefit and Privacy Panel (PBPP) for Health and Social Care. Local and regional data applications come under the consideration of the local Caldicott Guardianship governance structure.

Ethical and Health Research Authority (HRA) approval is a separate consideration and typically depends on the nature of the study. This is typically completed through the Integrated Research Application Service (IRAS). Decision tools regarding the requirement for ethical and HRA approval are available on the HRA website.[2] Evidence of appropriate Information Governance training is also required, typically in the form of completion of the Medical Research Council (MRC) Research, GDPR and confidentiality Quiz.[3]

Datasets

There are several databases available across the UK readily accessible for national data linkage projects. Details of those in England and Wales held centrally by NHS Digital can be found at: https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue,[4] whereas information regarding those held in Scotland can be found at: https://www.ndc.scot.nhs.uk/National-Datasets/index.asp.[5] For Northern Ireland, this is available here: https://www.nisra.gov.uk/support/research-support/administrative-data-research-northern-ireland-adr-ni.[6] There are also many other bespoke healthcare datasets that may be applicable to Trauma and Orthopaedics that can be identified through the Health Data Research UK (HDRUK) Innovation Gateway.[7] The HDRUK website is also an excellent resource for information on access to, and analysis of, health data, and in particular usage of 'Trusted Research Environments' that provide safe and streamlined access to healthcare information (see British Orthopaedic Association (BOA) research pages for a fuller version of this article, including a table of these datasets).

Researchers analyzing linked routinely collected healthcare data (such as those found in registries or national datasets) are encouraged to ensure that all design and reporting is compliant with the REporting of studies Conducted using Observational Routinely-collected Data (RECORD) statement.[8] Where feasible, results, data, and the code used for analyses should be published 'Open Access' to allow for the widest potential impact while maintaining transparency and reproducibility. This allows others the opportunity to reproduce analysis as well as learn from the methods used.

Once ethical and data access issues have been overcome, there are several issues that researchers should be aware of. Large routine datasets (such as HES) are cumbersome (consisting of many millions of rows of data each representing a hospital spell), take a lot of space to store, and as a result can take a long time to analyze. A simple calculation on several million rows of data may take hours to run, and as a result it is recommended that analysts create test or 'toy' datasets (small excerpts of the larger dataset) to ensure that the code works, before it is applied to the whole database.

Another important feature is to understand what data are missing, in what proportion, and the pattern of how they are missing. If data are missing in one group more than another, this can lead to biased analysis that may ultimately impact on the results and conclusions.

In summary, there are many opportunities available for data linkage projects across the spectrum of Trauma and Orthopaedics, with many excellent examples already published in high-impact literature. As the quantity of electronic healthcare information continues to grow exponentially, it is essential that these data are appropriately leveraged to perform high-quality research that helps to inform the future care of our patients, particularly as we move towards a more personalized and precise approach to healthcare.

Please see the following associated websites for more information: DARS;[9] EDRIS;[10] NISRA); and British Orthopaedic Association (BOA).[11,12]

Twitter

Follow L. Farrow @docfarrow
Follow J. Evans @jtevans13

References

1. **No authors listed**. Guide to the UK General Data Protection Regulation (UK GDPR): Lawful basis for processing. Information Commissioner's Office (ICO). 2023. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/ (date last accessed 9 March 2023).
2. **No authors listed**. What approvals and decisions do I need? NHS Health Research Authority. 2023. www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/ (date last accessed 9 March 2023).
3. **No authors listed**. Research, GDPR and confidentiality Quiz. UKRI Medical Research Council. 2023. https://byglearning.com/mrcrsc-lms/course/index.php?categoryid=1 (date last accessed 9 March 2023).
4. **No authors listed**. Data set catalogue. NHS Digital. 2023. https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue (date last accessed March 2023).
5. **No authors listed**. National Data Catalogue. ISD Scotland. 2023. www.ndc.scot.nhs.uk/National-Datasets/index.asp (date last accessed 9 March 2023).
6. **No authors listed**. Administrative Data Research Northern Ireland (ADR NI). Northern Ireland Statistics and Research Agency (NISRA). 2023. www.nisra.gov.uk/support/research-support/administrative-data-research-northern-ireland-adr-ni (date last accessed 9 March 2023).
7. **No authors listed**. Health Data Research Innovation Gateway. Health Data Research UK (HDRUK). 2023. www.hdruk.ac.uk/access-to-health-data/health-data-research-innovation-gateway/ (date last accessed 9 March 2023).
8. **Benchimol EI**, **Smeeth L**, **Guttmann A**, **et al**. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS medicine*. 2015;12(10):e1001885.
9. **No authors listed**. Digital Access Service Request (DARS). NHS Digital. 2023. https://digital.nhs.uk/services/data-access-request-service-dars (date last accessed 15 March 2023).
10. **No authors listed**. Electronic Data Research and Information Service (EDRIS). Public Health Scotland. 2023. https://www.isdscotland.org/products-and-services/edris/ (date last accessed 15 March 2023).
11. **No authors listed**. Research Support. Northern Ireland Statistics and Research Agency (NISRA). 2023. https://www.nisra.gov.uk/support/research-support (date last accessed 13 March 2023).
12. **No authors listed**. Research: Research Initiatives. British Orthopaedic Association (BOA). 2023. https://www.boa.ac.uk/research.html (date last accessed 15 March 2023).

Author information:
- L. Farrow, BSc, MBChB, MRCS, Clinical Research Fellow, University of Aberdeen College of Life Sciences and Medicine, Aberdeen, UK.
- J. Evans, BMedSci, MBBS, MSc, MD, FRCS, Hip Fellow, Trauma & Orthopaedics, Royal Devon & Exeter NHS Foundation Trust, Exeter, UK.

Author contributions:
- L. Farrow: Conceptualization, Writing – original draft, Writing – review & editing.
- J. Evans: Writing – review & editing.

Funding statement:
- The authors received no financial or material support for the research, authorship, and/or publication of this article.