**Findability of UK e-cohorts available for health research: mixed methods study Supplementary File 2**

**Discoverability of pre-specified health data resources**

Table shows whether a direct URL to the resource’s web site was found through the search engine searches, lists some of the indirect URLs found, and lists the catalogues in which the resources were listed. All links active as of 2018-08-07

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Search results** | |  |
| **Dataset name** | **Direct URL** | **Example indirect URLs** | **Catalogues** |
| **CPRD**  “Clinical Practice Research Datalink (CPRD) is a governmental, not-for-profit research service, jointly funded by the NHS National Institute for Health Research (NIHR) and the Medicines and Healthcare products Regulatory Agency (MHRA) , a part of the Department of Health and Social Care. We have been providing anonymised primary care records for public health research for thirty years.”  <https://www.cprd.com/home/> | <https://www.cprd.com/home/> | <https://www.herc.ox.ac.uk/downloads/health_datasets/browse-data-sets/general-practice-research-database-gprd>  <http://www.cancerresearchuk.org/funding-for-researchers/research-features/2017-07-05-using-patient-data-in-research-to-improve-cancer-treatments-and-increase-patient-survival>  <https://www.whitemoormedicalcentre-belper.co.uk/the-practice/your-health-records/>  <http://bmjopen.bmj.com/content/6/10/e012785>  <https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/NICE%20identifying%20data%202016%20Summary.pdf?ver=2016-10-18-120232-603>  <http://www.phgfoundation.org/documents/PHGF-linking-and-sharing-routine-health-data-for-research-in-england.pdf>  <https://acmedsci.ac.uk/file-download/35182-12569153801.pdf> | [www.hdf.nihr.ac.uk/datasets](http://www.hdf.nihr.ac.uk/datasets)  <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf>  <https://www.herc.ox.ac.uk/downloads/health_datasets/browse-data-sets> |
| **THIN**  “The Health Improvement Network (THIN) database represents a collaboration between two companies; In Practice Systems (INPS) - who developed Vision software used by general practitioners (GPs) in the UK to manage patient data, and IMS Health who then provide access to the data for use in medical research. THIN data are collected during routine practice and regularly delivered to THIN. Since THIN data collection began in 2003, over 500 Vision practices have joined the scheme.” <http://www.ucl.ac.uk/iehc/research/primary-care-and-population-health/research/thin-database/database> | not found | <https://www.ucl.ac.uk/pcph/research-groups-themes/thin-pub/database>\* LINK BROKEN  ***BING ONLY***  <http://www.ucl.ac.uk/iehc/research/primary-care-and-population-health/research/thin-database/database> LIVE LINK  <http://bmjopen.bmj.com/content/6/10/e012785>  <https://acmedsci.ac.uk/file-download/35182-12569153801.pdf>  <https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/NICE%20identifying%20data%202016%20Summary.pdf?ver=2016-10-18-120232-603> | <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf> |
| **QResearch**  “QResearch is a large consolidated database derived from the anonymised health records of over 30 million patients. The data currently come from approximately 1500 general practices using the EMIS clinical computer system. The practices are spread throughout the UK and include data from patients who are currently registered with the practices as well as historical patients who may have died or left. Historical records extend back to the early 1990's making it one of the largest and richest general practice databases in the world.”  <http://www.qresearch.org/> | <http://www.qresearch.org/> | <https://www.bsms.ac.uk/research/support-and-governance/governance-and-ethics/irp-ethics-and-governance/secondary-data-analysis.aspx>  <https://acmedsci.ac.uk/file-download/35182-12569153801.pdf> | <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf> |
| **ResearchOne**  “The research records contain linked data from a wide variety of health and care settings. ResearchOne maintains a comprehensive set of well structured data items, drawn from both administrative and clinical data across primary and secondary settings. This includes, for example, diagnostic codes, procedure codes, pathology test data, prescribing data, deprivation indices and care pathways. A full list of stored data items is available in the database summary or as part of a research data access application.”  <http://www.researchone.org/data/> | <http://www.researchone.org/data/> | <https://www.bsms.ac.uk/research/support-and-governance/governance-and-ethics/irp-ethics-and-governance/secondary-data-analysis.aspx>  <http://bmjopen.bmj.com/content/6/10/e012785>  <https://acmedsci.ac.uk/file-download/35182-12569153801.pdf> | <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf> |
| **CiPCA**  “The Consultations in Primary Care Archive (CiPCA) is a database of anonymised medical record data from a subset of general practices in North Staffordshire, UK.” <https://www.keele.ac.uk/mrr/cipcadatabase/> | not found | <https://www.keele.ac.uk/mrr/> |  |
| **HES**  “Hospital Episode Statistics (HES) is a database containing details of all admissions, A and E attendances and outpatient appointments at NHS hospitals in England.” <https://www.digital.nhs.uk/data-services/hospital-episode-statistics> | <https://www.digital.nhs.uk/data-services/hospital-episode-statistics> ***BING ONLY*** | <https://mrc.ukri.org/documents/pdf/obtaining-data-from-nhs-digital-v101016/>  <https://www.actuaries.org.uk/learn-and-develop/research-and-knowledge/guidance-researchers/access-data-sets>  <https://www.england.nhs.uk/wp-content/uploads/2013/07/hosp-data-consult.pdf>  <https://www.herc.ox.ac.uk/downloads/health_datasets/browse-data-sets/outpatients-hospital-episode-statistics-hes>  content.digital.nhs.uk/datasets | <http://www.hdf.nihr.ac.uk/datasets>  <https://www.herc.ox.ac.uk/downloads/health_datasets/browse-data-sets> |
| **SAIL Databank**  “The SAIL Databank is a world-class complete solution to sourcing, accessing, linking and analysing health and population data all within a governed infrastructure that is safe and secure. Researchers can access a broad range of routinely collected data spanning up to 20 years from an entire population.”  <https://saildatabank.com/> | <https://saildatabank.com/> | <https://adrn.ac.uk/media/174239/admin_health_data_in_sail-data_brief-web.pdf>  <https://journals.edizioniseed.it/index.php/FE/article/viewFile/677/770>  <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-9-3>  <https://ijpds.org/article/view/286> | <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf> |
| **SIR**  “In 2007, the Salford Integrated Record (SIR) was established with real time electronic links between primary and secondary care. The SIR linked these sources with patients’ prescription information, capturing data on everyone accessing health services in Salford. In 2008, the NorthWestEHealth (NWEH) group was established to use SIR data for research. The group brought together computer scientists, clinicians, researchers, statisticians, epidemiologists, and project managers to develop methods and a governance framework to extract patient healthcare data for use in real world research studies.”  <https://nweh.co.uk/news/2018/putting-patients-in-control-of-data-from-electronic-health-records> | not found | <https://nweh.co.uk/news/2018/putting-patients-in-control-of-data-from-electronic-health-records> ***GOOGLE ONLY*** |  |
| **PIS (Scotland)**  “The Prescribing Information System (PIS) is the definitive data source for all prescribing relating to all medicines and their costs that are prescribed and dispensed in the community in Scotland. The information is supplied by Practitioner & Counter Fraud Services Division (P&CFS) who is responsible for the processing and pricing of all prescriptions dispensed in Scotland. These data are augmented with information on prescriptions written in Scotland that were dispensed elsewhere in the United Kingdom.”  <http://www.ndc.scot.nhs.uk/National-Datasets/data.asp?SubID=9> | not found | <http://www.adls.ac.uk/wp-content/uploads/Introduction-to-ISD-administrative-data.pdf> LINK BROKEN, no live ***GOOGLE ONLY*** | <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf>  <http://www.ndc.scot.nhs.uk/National-Datasets/index.asp> |
| **NHS Lothian Research Safe Haven / The University of Edinburgh**  “The Research Data Linkage Service works in partnership with the NHS Lothian Research Safe Haven, which is one of a network of regional safe havens in Scotland, providing record linkage of routinely held Lothian data for the purpose of furthering research.”  <https://www.ed.ac.uk/usher/edinburgh-clinical-trials/data-linkage> | <https://www.ed.ac.uk/usher/edinburgh-clinical-trials/data-linkage> ***GOOGLE ONLY*** | <https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens> broken  <http://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens-2> live  <https://academic.oup.com/bjsw/article/45/5/1614/1629133> |  |
| **Grampian Data Safe Haven**  “NHS Grampian Routine Health Data Grampian Data Safe Haven (DaSH) acts as a data processor and interfaces with NHS Grampian Health Intelligence who provide access to routinely collected NHSG health care data for data linkage research.”  <https://www.abdn.ac.uk/iahs/research/health-informatics/datasets.php> | not found | <https://www.abdn.ac.uk/iahs/research/health-informatics/datasets.php> ***GOOGLE ONLY***  <https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens> broken  <http://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens-2> live |  |
| **Health Informatics Centre - Trusted Research Environment (University of Dundee)**  “HIC hosts a range of linkable routinely-collected, person-based administrative datasets combining datasets for the Tayside and Fife population and Tayside, with local speciality research, and clinical datasets.‌‌”  <https://www.dundee.ac.uk/hic/datalinkageservice/datasetinventory/> | <https://www.dundee.ac.uk/hic/datalinkageservice/datasetinventory/> | <https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens> broken  <http://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens-2> live | <http://www.farrinstitute.org/wp-content/uploads/2017/10/Datasets-that-may-be-of-interest-to-Primary-Care-Researchers-in-the-UK-May-2016.pdf> |
| **NHS Greater Glasgow & Clyde Safe Haven**  “We form part of the Scotland branch of The Farr Institute, which is a UK-wide network of federated Safe Havens. We are the result of a unique collaboration between NHS Greater Glasgow & Clyde Research & Development and the Robertson Centre for Biostatistics, offering an excellent resource for health-related data research in Scotland.”  <http://www.nhsggc.org.uk/about-us/professional-support-sites/nhsggc-safe-haven/> | <http://www.nhsggc.org.uk/about-us/professional-support-sites/nhsggc-safe-haven/your-research/available-datasets/> | <https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens> broken  <http://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens-2> live |  |