|  |
| --- |
| **Data Portal**  Work Package 2 |
| **Objective(s)**: |
| This work package was established to facilitate the integration of cohort data and UK wide data linkage into a DPUK Informatics Portal. It has involved the collection, integration, curation and analysis of DPUK data.  Four specific objectives were set initially:   1. Understanding the data capability of UK wide record linkage 2. Integration of cohort research data and metadata 3. Establishment of the Informatics Portal 4. Dissemination aligned with conveying the capabilities and potential of the Informatics Portal   As this work is fundamental to DPUK, this work package has been subject to constant change and evolution. The work undertaken, while still falling within the four objectives above, has been significantly expanded beyond the original objectives. |
| **Overview Summary:** |
| A fundamental objective of Dementias Platform UK was the establishment of an Informatics Portal. It was to be capable of integrating data related to cohorts of interest and to be able to accommodate data linkage. The DPUK Data Portal (<https://portal.dementiasplatform.uk>) is the result of this and it is continually being refreshed and enhanced in line with feedback from users. The Portal itself as attracted international research attention and has created a network of 42 partner cohorts, plus independent research collaborators, who choose to store and share data using the Data Portal and also facilitate their own data sharing projects using the Portal. |
| **Executive Summary:** |
| In the original bid DPUK submitted to MRC, the DPUK Data Portal was designed to support epidemiological studies into dementia utilising UK cohorts. It has been successful, overcoming significant reluctance from some cohort PIs and institutional lawyers to providing data to the wider scientific community. The DPUK Data Portal provides access to more cohorts and participant data than any other system across the globe (over 3 million people) in a free-to-access resource. Continued involvement with the research community and feedback from the MRC Oversight Board has led to considerable changes in design and functionality, all designed to enhance the user experience and facilitate multi-modal data analyses. It currently supports complex statistical analysis, machine learning studies, small-scale genomic analyses and some imaging analysis, all in one analysis platform. DPUK2 will further enhance these offerings to maximise the potential of the integration and product development begun in DPUK1. |
| **Summary of Outputs**: (as per Researchfish categories) |
| **Publications:** |
| * Orton C., et al. (2019) **Dementias Platform UK Data Portal – Multi-Modal Data Analysis Using Customised Infrastructure.** Alzheimer’s and Dementia, 15(7), p474 * Orton C., et al. (2018) **Dementias Platform UK (DPUK) Data Portal – world-leading infrastructure facilitating innovative multi-modal research.** International Journal of Population Data Science, 3(4) * [Orton C., et al. (2018).](https://app.researchfish.com/portfolio/0/publications/5c87c11e5fe508.15840101/view?name=5c87c11e5fe508.15840101&delegator=39791&filter=MRC--MR/L023784/2) **Dementias Platform UK (DPUK) Data Portal - supporting multi-modal data analysis, data linkage and real-world outcomes.** International Journal of Population Data Science, 3(2) * [North L., et al. (2018).](https://app.researchfish.com/portfolio/0/publications/5c87c11e3e55a8.12586939/view?name=5c87c11e3e55a8.12586939&delegator=39791&filter=MRC--MR/L023784/2) **Data Services for Cohort Studies: Increasing the impact of existing research studies and epidemiological readiness.** International Journal of Population Data Science, 3(4) * [Gallacher J, (2016).](https://app.researchfish.com/portfolio/0/publications/5aa95927efe497.75263262/view?name=5aa95927efe497.75263262&delegator=39791&filter=MRC--MR/L023784/2) **What's the good of a science platform?** Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 374 (2083), * Bauermeister S. et al. **The Dementias Platform UK (DPUK) Data Portal (2020)** Eur J Epidemiology   <https://doi.org/10.1007/s10654-020-00633-4> |
| **Collaborations & Partnerships** |
| By its nature the Data Portal is a collaborative project with significant partnerships established to meet objectives and facilitate development.  Alzheimer’s Disease Data Initiative – Gates Ventures, CPAD, GAAIN, HDR UK Health Data Research UK Datathon programme – Alzheimer’s Research UK, Turing Institute |
| **Further Funding** |
| The Data Portal has been supported as part of renewed funding for DPUK that will be available from 2021. |
| **Next Destinations** |
| Chris Orton – previously Data Project Manager for the Data Portal, has been promoted to Programme Manager within Population Data Science at Swansea University, now leading on the BREATHE Health Data Research Hub for Respiratory Health, and other cross-cutting initiatives.  Laura North – previously data analyst working on cohort data curation and Data Portal data discovery tools, has been promoted to a Research Officer and Data Scientist post within the Wales and Northern Ireland HDR UK site at Swansea University.  Mark Newbury – Mark is now the lead contact for the Data Portal having taken on increased responsibilities. |
| **Engagement Activities** |
| ***2020***   * ADDI Project Working Group – Miami, February 2020 * HDR UK Scientific Strategy Board – January 2020   ***2019***   * HDR UK Scientific Strategy Board – November 2020 * DPUK Datathon – November 2019, Swansea * DPUK Datathon – October 2019, UEA * ADDI Project Working Group – London, October 2019 * ADDI Project Working Group – Los Angeles, July 2019 * Orton C. AAIC 2019 – Portal poster presentation, Los Angeles, July 2019 * DPUK Datathon – May 2019, Exeter * Orton C. – IASIS: Big Data for Precision Medicine, Plenary Meeting, March 2019 * Orton C. – EU Joint Programme - Neurodegenerative Disease Research, ADDITION Consortium Kick-off meeting, regarding the DPUK Data Portal supporting the research analysis, February 2019 * Orton C. – Health Data Research UK Site Meeting - January 2019   ***2018***   * Bauermeister S – Oxford Dementia Research Day 2018 Poster: The MRC DPUK Data Portal: Bringing Researchers to Data: The Research Journey: * Bauermeister S – DPUK Data Portal Workshop Presentation at the DPUK Annual Conference * IPDLN 2018 DPUK Data Portal presentations (Chris Orton + Laura North) * DPUK Datathon – October 2018, Turing Institute- 42 registered participants and users * North L. – International Population Data Linkage Network Conference - Banff, Canada, September 2018 - Data Services presentation 2018 * Bauermeister S - ARUK Technical Workshop * Orton C. – Global Alzheimer's Disease Data Innovative - Interoperability Working Group, Toronto, December 2018 * Orton C. – Nature Neuroscience Interoperability Meeting, Toronto, December 2018 - Data Portal presentation and panel appearance * Orton C. – International Population Data Linkage Network Conference - Banff, Canada, September 2018 - Data Services presentation * Orton C. – Administrative Data Research Network Conference - Belfast, June 2018 * Orton C. – NHS Digital Collaboration Meeting * Orton C. – DPUK Annual Conference - Data Portal presentation session * Orton C. – Alzheimer's Research UK Conference - London - Poster Presentation * Orton C. – DPUK Cohort Workshop * Orton C. – Brains for Dementia Research Study Day * Orton C. – Member of the governance, design and delivery strategy groups for the MS Society * Gallacher, J. – Participant at the International Cohorts Summit * Gallacher, J. – DPUK MRI-PET Imaging Network Launch * Gallacher, J. – Lecturing as part of Oxford University's MSc in Non-communicable diseases - neurodegenerative conditions with a focus on dementia * Gallacher, J. – Talk at a workshop with AMED on Ageing Society * DPUK Turing Institute Datathon 2018 DFP Cohort * Lyons R – 2/10/18 - Technology Sector Symposium - Developing a UK Wide Health Data Research Platform * Lyons R – 5/9/18 - Design Thinking Workshop * Lyons R – 29-30/5/18 - International Clinical Phonemics’ Workshop * Lyons R *–* 1 /5/18 - Swansea University HDR UK & NHS Digital Data Workshop   ***2017***   * Orton C. – Presentation at Database Resources for Large-Scale and Complex Data workshop, Engineers' House, Bristol, 21st June 2017 * Orton C. – Presentation on DPUK as a use case of UKSeRP at the Administrative Data Research Network Conference, Royal Surgeon's Hall, Edinburgh, 2017 * Orton C. – DPUK Data Portal Live Demonstration, Swansea University, at the DPUK Annual Scientific Meeting 2017 * Thompson S. – DPUK Infrastructure Presentation at the EMIF DPUK Joint Conference, February 2017 * Bauermeister S – Poster presentation: The MRC DPUK Data Portal: Bringing Researchers to Data: The Research Journey: Alzheimer's Research UK Conference 2017 * DPUK infrastructure twice presented at ADRN Conference 2017 2nd June * DPUK Attendance as an exhibitor and featured session at AAIC 2017 19th July * Orton C. – Anti-Microbial Resistance Databank collaboration meeting * Orton C. – AAIC 2017 - Dementias Platform UK Showcase presentation * Lyons R – Presentation to MRC Council - Centre for the Improvement of Population Health through E-records Research (CIPHER) *2016* * Orton C. – DPUK Cohort Collaboration Meeting * Orton C. – DPUK Presentation at the International Population Data Linkage Network Conference, Swansea, August 2016 * Orton C. – DPUK Cohort Seminar * Orton C. – Outline presentation of work package 2 of DPUK, the informatics stream, at the time focusing on the build and purpose of the DPUK Data Portal. The poster featured a description of the envisaged tools, data management and planned activity for the coming year.   Lyons R – Big Data – next big thing: evaluation. Presentation to 1st NHS Wales Digital Health and Care Conference |
| **Influence of policy, practice, patients & the public** |
| Expedited legal review of data sharing agreements following precedent-setting system set up by Dementias Platform UK. |
| **Research Tools & Methods** |
| DPUK Data Portal Analysis Environment  DPUK Data Portal website and tools – [https://portal.dementiasplatform.uk](https://portal.dementiasplatform.uk/) |
| **Research Databases & Models** |
| DPUK Data Portal: Data Resource Profile –<https://www.biorxiv.org/content/10.1101/582155v1>  DPUK Data Portal and DPUK contribution to the following dataset (SAIL Dementia eCohort):<https://ijpds.org/article/view/1121> |
| **Intellectual property & licencing** |
| None |
| **Medical products, interventions & clinical trials** |
| None |
| **Artistic & creative products** |
| None |
| **Software & technical products** |
| None |
| **Spin outs** |
| None |
| **Awards & recognition** |
| None |
| **Other outputs & knowledge/future steps** |
| DPUK2 will build upon the initial success of DPUK1 to further develop its systems ensuring market-leading capability and capacity for dementias and neuroscience research. Particular focus will be made on integrating complex -omics and imaging data into a centralised research platform, as well as further developing the existing data discovery tools. |
| **Use of facilities & resources** |
| Use of MRC funds to design, build, and establish the DPUK Data Portal and integrated systems such as the imaging network and federated query capability for the ADDI programme. |
| **Most successful outcome and what it means for future dementia research**: |
| DPUK provides a technical platform which allows for a variety of virtual desktop sizes supporting multi-modal research into Dementia and neuroscience. To date, this has supported 106 studies. The platform is increasingly supported by a range of tools to help researchers explore metadata prior to application and continues to grow in popularity. By providing a central platform for dementia research, a collaborative space for research has emerged and has provided a service for data providers and researchers alike to more conveniently share and access data, as well as providing support and a platform for a more interlinked dementia research community making use of rich cohort datasets. |
| **Lessons learned**: |
| * Cohorts are often overly busy and understaffed. It is necessary to frequently chase and cohorts generally appreciate reminders of pending applications via email. It is necessary to take a strongly proactive approach to contacting cohorts regarding decisions, and, where necessary, data upload in order to expedite access to the data for researchers. * Cohort data is often never provided in full, particularly where cohort are continuing to collect data. It is therefore important to ensure users are aware of the limitations to data availability but also encourage cohorts to expedite data sharing which is required under the conditions of the funding for the data collection. * Deliverables should mirror the changing landscape of dementia research to enable refocusing of resources and delivery aims as necessary. The Data Portal for instance was created with the aim of being purely an epidemiology platform, and now houses   -omics and imaging capability in one system, rather than in a federated network, which came about organically. Deliverables, therefore, although will be overarching to begin with, need to adapt and be amended to remain relevant throughout the project.   * Governance in general is a strong part of the DPUK Data Portal and DPUK is setting standards in this area. However, it is notable that cohorts, and particularly data sharing are still subject to long-winded legal proceedings. The services offered by the Data Portal should be promoted more to show that the Portal benefits not only the third-party research community but cohorts themselves. |
| **Other:** |
| None |
| **Date of Report**: |
| 11 September 2020 |