## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>01</td>
</tr>
<tr>
<td>Research Highlights</td>
<td>03</td>
</tr>
<tr>
<td>- The Farr Institute’s Cross-Centre Projects</td>
<td>03</td>
</tr>
<tr>
<td>- Exemplars of The Farr Institute’s Research 2016-17</td>
<td>05</td>
</tr>
<tr>
<td>Enabling Research at Scale: datasets, infrastructure, methods, standards, governance</td>
<td>09</td>
</tr>
<tr>
<td>- Trusted Research Environments</td>
<td>09</td>
</tr>
<tr>
<td>- Discovery Across Data Sectors</td>
<td>11</td>
</tr>
<tr>
<td>- Analytic Methodology</td>
<td>11</td>
</tr>
<tr>
<td>- New Data Feeds</td>
<td>11</td>
</tr>
<tr>
<td>- Actionable Analytics</td>
<td>13</td>
</tr>
<tr>
<td>Capacity Building</td>
<td>14</td>
</tr>
<tr>
<td>- The Farr Institute Doctoral Training Programme</td>
<td>14</td>
</tr>
<tr>
<td>- Continuing Professional Development</td>
<td>15</td>
</tr>
<tr>
<td>- NHS Collaborations: Modernising Scientific Careers &amp; NHS Digital Academy</td>
<td>15</td>
</tr>
<tr>
<td>- The Farr Institute Future Leaders in Health Data Science</td>
<td>15</td>
</tr>
<tr>
<td>- Future Plans: Early Career Researcher Exchange Programme</td>
<td>16</td>
</tr>
<tr>
<td>Innovative Governance</td>
<td>17</td>
</tr>
<tr>
<td>- Influence</td>
<td>17</td>
</tr>
<tr>
<td>- Tools and Solutions to Enhance Data Sharing and Use</td>
<td>17</td>
</tr>
<tr>
<td>- Engagement</td>
<td>18</td>
</tr>
<tr>
<td>- Dissemination</td>
<td>18</td>
</tr>
<tr>
<td>Public and Patient Involvement &amp; Engagement</td>
<td>19</td>
</tr>
<tr>
<td>- Public Panel Activity</td>
<td>19</td>
</tr>
<tr>
<td>- Citizen Driven Health</td>
<td>19</td>
</tr>
<tr>
<td>- Resources for the Publics</td>
<td>20</td>
</tr>
<tr>
<td>- Website and Social Media</td>
<td>21</td>
</tr>
<tr>
<td>- Events and Activities</td>
<td>22</td>
</tr>
<tr>
<td>- Mainstream Press and Media Coverage</td>
<td>23</td>
</tr>
<tr>
<td>Partnerships</td>
<td>24</td>
</tr>
<tr>
<td>- Partnering with Industry</td>
<td>24</td>
</tr>
<tr>
<td>- Multi-partner Large Scale Collaborations</td>
<td>25</td>
</tr>
<tr>
<td>- Formalising a Partnership with the Institute of Clinical Evaluative Sciences, Ontario (ICES)</td>
<td>26</td>
</tr>
<tr>
<td>- Partnering to Expand the Reach of The Farr Institute International Conference</td>
<td>26</td>
</tr>
<tr>
<td>Appendices</td>
<td>28</td>
</tr>
<tr>
<td>- Appendix 1: Publications</td>
<td>29</td>
</tr>
<tr>
<td>- Appendix 2: Markers of Esteem</td>
<td>40</td>
</tr>
<tr>
<td>- Appendix 3: Leveraged New Grant Funding</td>
<td>46</td>
</tr>
<tr>
<td>- Appendix 4: Patient and Public Involvement and Engagement activities</td>
<td>50</td>
</tr>
</tbody>
</table>
Executive Summary

This past year The Farr Institute has continued to deliver on its mission to position the UK as a world leader in health informatics research through scientific discovery and the enhancement of patient and public health. Across the 21 academic institutions that make up the Institute, Farr researchers have published an average of six peer-reviewed papers every week over the course of the year.

Our goal in 2016/17 has been to focus our activities across six key domains that act as our guiding principles and provide strategic steer [Box 1]. Most importantly, following advice from our International Advisory Board and funders, we have with renewed determination forged collaborations across the UK so that we act as a single Institute by convening stakeholders from different sectors and localities to share information, establish best practices, enable interoperability and provide expertise and know-how in project management, ethics, regulation, public engagement, technology development and co-ordination of governance. We have also initiated several research programmes that aim to demonstrate the ability to catalyse the sharing and analysis of data at scale. We hope this shared learning will help provide a solid foundation for the new UK Institute for health and biomedical informatics research, Health Data Research UK (HDR UK).

Research excellence is a primary goal, and this year, Farr researchers have published 238 papers, of which 36 were in papers with an impact factor greater than eight. We have also provided pump-prime funding for six high impact cross-Centre research programmes that will facilitate research at scale, and answer questions that individual Farr Investigators would be unable to pursue in isolation, many with collaborators from other organisations beyond healthcare research. It was also a year of leverage with The Farr Institute being the catalyst for several large programme grants including the €20 million Innovative Medicines Initiative (IMI) cardiovascular award (Farr London, Scotland, CIPHER) and £2 million Innovative Healthcare Delivery Programme on Cancer (Farr Scotland).

This year we launched several initiatives to demonstrate secure inter-operability between research data infrastructure investments across the UK, to facilitate data sharing and federated meta-analyses. All Farr Centres are actively participating in the Jisc Safe share project that is at the cutting edge of providing this functionality for health research. The Farr Institute UK Secure eResearch Platform (UKSeRP), continues to go from strength to strength, providing underpinning technology for Dementia Platform UK, with the enticing prospect of hosting up to 33 dementia related cohort studies for collaborative research. The Farr Institute in Manchester has provided the research and analytical platform for Connected Health Cities, a cornerstone of the Northern Powerhouse. Likewise, our quest to mobilise new datasets for research gains momentum, with notable developments in national prescribing and national imaging in Scotland and incorporation of local authority and other public datasets in Wales. Partnerships with The Alan Turing Institute and multiple industry partners are providing novel insights into health data science.

Capacity building is another important goal. This past year we have co-hosted numerous events to create a thriving community of doctoral and post-doctoral researchers across the UK, including over 100 PhD students and the successful launch of The Farr Institute Future Leaders programme, with the aim of training the next generation of fearless leaders in health informatics. This has been supplemented with co-ordinated Masters, CPD and MOOC provision, reaching over 5000 people.

Ensuring privacy, security good governance and public engagement is paramount. In Wales and Scotland, The Farr Institute has continued to build good relationships with NHS, Government and local authority data controllers and this has been instrumental to facilitate access to data, proportionate governance and improve data linkage. In Scotland, The Farr Institute co-designed the Public Benefits and Privacy Panel, launched in 2016, which is a single place of scrutiny for linked data applications. Over 250 projects have now used The Farr Institute technology and governance.

---

Box 1. The Farr Institute Objectives, from the Annual Report 2015/16

1. Perform pioneering multi-disciplinary research with large and complex health-related data, including healthcare records and biological, social, and environmental data.

2. Enable new datasets and develop new infrastructure, methods, technologies, and standards for such research.

3. Develop skills, talent and expertise in individuals and research communities for collaborative working.

4. Work with the owners and controllers of data to support the safe use of patient information for medical research across the UK, championing data protection, confidentiality and privacy.

5. Engage with the public to demonstrate the benefits of using health data in research and to encourage the support of secure and trusted access to patient information.

6. Bring together government, public sector, academia and industry to foster relationships and establish best practices for innovation, discovery and impact in health data science.
infrastructure in Scotland. In England, there is senior Farr Institute representation on the research advisory group for NHS Digital, and we have made important contributions to the National Data Guardian (NDG) Consultation, Industrial Strategy and support for Jo Churchill MP’s NDG bill.

Finally, we have had a productive year with public engagement, with multiple exposure in science festivals, creation of the 100 Ways case study series, further development of the #datasaveslives campaign, and strategic input into the Understanding Patient Data initiative led by Wellcome Trust. Many of these materials and activities have been widely used by many organisations and individuals nationally and on social media. Our online presence continues to grow with over 24,000 users of the new Farr Institute website in the past year. All Centres have continued to increase the amount of research co-developed with members of the public.

Since the launch of The Farr Institute, the challenges of leading a single national research Institute have been varied and challenging. This past year was no exception. 2016/17 was notable for the tremendous productivity of our staff and scientists, numerous high impact publications and, for the first time, the ability to execute research and data sharing across the UK. We believe we have strengthened partnerships locally, nationally and internationally, but there is still much to be achieved. Misconceptions about the original purpose of The Farr Institute remain, and the challenges of acting as a single research site across complex data systems are under-estimated by many. Harmonisation of data access, governance and technology and data standards remains challenging, especially in England. We are proud however to have built a vibrant, forward looking community in health informatics as demonstrated by the over 750 registrants for the Informatics for Health 2017 Farr Institute conference in Manchester in April 2017, compared with 120 registrants in 2013.

In 2017/18, we seek to support five activities; (i) continue to deliver on The Farr Institute six priorities as part of the original eHIRC grants; (ii) place a greater emphasis on the scale of research across the UK; (iii) link with bio-informaticians and experimental medicine researchers to harness structured, unstructured and “omic” data for discovery science and precision medicine; (iv) further develop links with The Alan Turing Institute and industry to build collaborative research programmes in health data science; and (v) support the successful launch of Health Data Research UK.

We would like to take this opportunity to thank The Farr Institute teams across the UK for their relentless enthusiasm and support, as well as our collaborators and partners in NHS, Government, other public sector bodies and industry. We would like to thank our funders and also our International Advisory Board, chaired by Professor Nancy Pedersen, for their vision and support. Together we will pursue our quest to a world-leading institute where data, discovery and partnership improve health and healthcare.

Professor Iain Buchan
Iain Buchan is Clinical Professor in Public Health Informatics at the University of Manchester, where he founded and leads the Centre for Health Informatics.

Professor Harry Hemingway
Harry Hemingway is Professor of Clinical Epidemiology at University College London and Director of the Institute for Health Informatics, UCL.

Professor Ronan Lyons
Ronan Lyons is Professor of Public Health, Swansea University and Honorary Consultant in Public Health with Public Health Wales.

Professor Andrew Morris
Andrew Morris is Professor of Medicine, Director of the Usher Institute of Population Health Sciences and Informatics and Vice Principal of Data Science at the University of Edinburgh.
Research Highlights

Objective 1; Perform pioneering multi-disciplinary research with large and complex health-related data including healthcare records and biological, social, and environmental data

The Farr Institute has continued to deliver impactful, pioneering research in 2016/17 building on the funding foundations of the original eHIRC grants. As described in the annual report 2015/16 we have focussed on research strengths in themes such as methodological development, data analytics, maternal and child health, cardiovascular disease, asthma, kidney disease, cancer, dementia, primary care, public health and new data streams. We have continued to expand our discovery research in combining ‘omics and electronic health records through strategic partnerships with Genomics England, MRC Biomedical Centres, NIHR Biomedical Research Centres and UK Biobank. University College London and University of Manchester were successful in retaining (UCL) or achieving (Manchester) NIHR BRC funding in 2016, both with a strong informatics theme. Farr Scotland is part of both the ‘Precision Medicine Ecosystem’, a £4 million funding award from Scottish Government, and a partnership with AstraZeneca and NHS Scotland to apply genomic knowledge to the development of innovative medicines and targeted treatments.

We have also sought to scale research initiated within one Farr Centre across the UK and have used Network funds to invest in small projects which either built on existing research, for example expanding the size of the study population or allowing research to be repeated in a comparator study group, or allowed pilot studies to be initiated to explore the feasibility of the research and provide data which could be used as the basis to apply for further funding. The projects were also designed to allow for the exploration of the challenges faced by researchers trying to collaborate across the UK, for example access to linked data held by different data controllers, to provide real-world situations which could then be solved to facilitate similar research in the future. A number of these projects are summarised on pages 3-4.

2016/17 saw further ‘Frontiers Meetings’ bringing together stakeholders across the UK to develop new collaborative research. Subjects included: substance misuse; environmental determinants of child health and cognitive function; risk prediction and pregnancy.

The Farr Centres have leveraged over £83 million of new funding (16% RCUK, 6% Industry, 13% Charities, 34% NIHR, 8% International) in 2016/17.

The investment from the funders in The Farr Institute has directly or indirectly supported research leading to 238 publications in 2016/17. 36 of these publications are in journals with an impact factor greater than eight. The complete list of publications can be found in appendix 1. On pages 5-8 we have selected eight published research papers which exemplify the impact The Farr Institute’s research has had from discovery science to public health.

The Farr Institute’s Cross-Centre Projects

One of the ambitions of The Farr Institute is to carry out research at scale across the UK. This is more challenging than it seems due to the patchwork quilt of data availability in different jurisdictions and regions and, as yet, unfinished work in combining data across accessible platforms. In essence, each area represents an incomplete and partially overlapping jigsaw puzzle of data assets. Finding the exact same data in the same format in more than one location can be surprisingly difficult. Even then, slightly different source definitions and data structures need to be investigated and harmonised before analysis can be combined. Below are a number of exemplar Farr Institute projects that use data from more than one area of the UK.

1. Health outcomes of children and adults with learning difficulties

This project seeks to investigate factors associated with the diagnosis, treatment, health and educational outcomes of vulnerable groups in society, particularly those with learning difficulties and/or Autistic Spectrum Disorders. A multi-disciplinary team from the Universities of Glasgow, Cardiff and Swansea with expertise in epidemiology, learning disorders and health informatics have been creating comparative cohorts from multiple sources.
Research Highlights

including education, general practice, inpatient, mortality and population registers. The population cohorts involve more than 1.5 million linked records and the first analysis on relative risk of hospitalization from all and selected causes has commenced.

2. Temporal trends in outcomes following diagnosis of Down’s Syndrome

Both Scotland and Wales operate high quality registers of children born with congenital anomalies. There have been considerable advances over the past 20 years in the provision of care for children born with complex syndromes. This study aims to quantify changes in mortality and morbidity over that time period by creating population based e-cohorts from routine data in Wales and Scotland to study all-cause mortality, cause-specific mortality, hospital admissions (disease and procedure codes) and medication use (cardiac, diabetic, antidepressants). Funding from The Farr Institute has been used to create a Scottish cohort to compare with data from the Wales Electronic Cohort for Children.

3. The risk of death and all serious vascular events for patients who start a non-vitamin K Oral Anticoagulant after intracranial haemorrhage

A new class of oral anticoagulant drugs, Target-Specific Oral Anticoagulants (TSOACs), has been introduced with evidence of a favourable risk-benefit profile. However, there is some concern that the current lack of a TSOAC antidote creates a significant risk different to the traditional anticoagulant treatment, warfarin. This study, which includes researchers from the Universities of Edinburgh, Swansea and Athesina Studorum (University of Trento, Italy) aims to clarify this uncertainty by evaluating the relative risks of death and serious vascular event risks in those treated with TSOACs versus warfarin, following intracranial haemorrhage in Scottish, Welsh and Italian e-cohorts. Data from multiple countries is required to produce precise estimates of risk. Our intention is to deliver public benefit from better understanding any risks associated with prescription of the new TSOACs to specific groups leading to greater public safety. Researchers have had joint access to both the Scottish eDRIS and Welsh SAIL Databank safe havens, and work has begun on the Welsh population to complete a validation against the previously completed project based on the Scottish population.

4. Improving the efficiency and transparency of pharmacoepidemiology through standardisation of data preparation

Anonymised electronic patient records regularly support pharmacoepidemiological studies, providing real-world evidence about the comparative benefits and harms of drug treatments. Prior to analysis, data preparation is required to transform ‘messy’ prescription data into formatted ‘clean’ drug exposure data. At present, research groups each devise their own pathway for data preparation, with differing assumptions, and do not share their methods. No framework for conducting or reporting data preparation currently exists. This project aims to develop an algorithm which would allow the re-use of methods, thereby improving the efficiency of pharmacoepidemiology. This aligns with EQUATOR’s REWARD initiative and would benefit both researchers and funders. It would also provide a reporting structure for the recent RECORD statement (www.record-statement.org). Work trialling the algorithm on different datasets is underway and the groups involved will be reporting their results in March 2017.

5. Pharmacoepidemiological surveillance for children with chronic conditions using data mining

The project had two aims; (1) To demonstrate the use of machine learning for pharmacoepidemiology surveillance of medications used in childhood, using 2 exemplar cases and (2) to submit a grant to develop a pharmacoepidemiology platform. Manuscripts based on the development of the machine learning methodology for surveillance and the application of the machine learning tools to monitor the use of antipsychotic medication are in preparation. The work has also generated a dataset on the pregnancy outcomes of girls who were prescribed antipsychotic medication in teenage years which is also being analysed ready for publication. The data generated is being used to support funding applications to the British Heart Foundation and the Wellcome Trust.

6. Identifying episodes of acute kidney injury (AKI) across healthcare settings using routinely collected data

The Kidney Disease@Farr collaboration brings together a multidisciplinary group of researchers with interests in health informatics and kidney disease. This project aimed to evaluate and improve the portability of an existing AKI electronic phenotyping algorithm developed in Grampian, Scotland. The algorithm has now been implemented across three other UK regions: Hampshire, Salford and Swansea. All datasets consisted of linked electronic health record data across primary and secondary care. Using the algorithm, researchers compared the incidence and character of AKI episodes in each region and index year and found substantial differences between regions in AKI incidence. The group are currently systematically investigating potential reasons for these differences. Anticipated outcomes of the project in the next few months are: a publically available AKI phenotyping algorithm, including associated metadata and guidance for local implementation; a clinically focussed paper on incidence and characteristics of AKI episodes in four UK healthcare populations, including potential explanations for differences; a methodological paper providing guidance on how to implement electronic phenotyping algorithms across datasets that have different contexts and underlying infrastructures.

Illustration of Andrew Morris’ talk, BioData World Congress 2016.
Copyright Two Viz Think Limited.

www.farrinstitute.org 4
Exemplars of The Farr Institute’s Research 2016/17

1. Richer Understanding of Unmet Need: The epidemiology, healthcare and societal burden and costs of asthma in the UK

The Challenge
There are a lack of reliable data on the epidemiology and associated burden and costs of asthma. This study, done in collaboration with the UK Asthma Centre for Applied Research, and across Farr Centres, sought to provide the first UK-wide estimates of the epidemiology, healthcare utilisation and costs of asthma.

The Research
The study found that asthma resulted in at least 6.3 million primary care consultations, 93,000 hospital in-patient episodes, 1,800 intensive-care unit episodes and 36,800 disability living allowance claims in the UK during 2011/12.

The total costs of asthma for this period were estimated at least £1.1 billion. 74% of these costs were for provision of primary care services, 13% for disability claims, and 12% for hospital care. There were 1,160 asthma deaths.

The prevalence of asthma depended on the definition and data source used. The UK lifetime prevalence of patient-reported symptoms suggestive of asthma was 29.5% (95% CI, 27.7-31.3; n = 18.5 million people) and 15.6% (14.3-16.9; n = 9.8 million) for patient-reported clinician-diagnosed asthma. The annual prevalence of patient-reported clinician-diagnosed-and-treated asthma was 9.6% (8.9-10.3; n = 6.0 million) and of clinician-reported, diagnosed-and-treated asthma 5.7% (5.7-5.7; n = 3.6 million).

This study clearly illustrates that asthma is very common and is responsible for considerable morbidity, healthcare utilisation and financial costs to the UK. Since much of the morbidity and mortality is considered potentially preventable, greater policy focus is needed to reduce the risk of asthma exacerbations, hospitalisations and deaths, and reduce costs.

The Impact
This study provided the first UK-wide estimates of the epidemiology, healthcare utilisation and costs of asthma. The total costs of asthma in the UK during 2011/12 were estimated at least £1.1 billion.

The UK lifetime prevalence differed with definition and data source, varying between 3.6 million and 18.5 million people.

Reference

2. Discovering Drugs Better: The druggable genome and support for target identification and validation in drug development

The Challenge
Drug development has a low probability of success with only 4% of drug development programmes yielding licensed drugs. This can largely be attributed to unresolved system flaws whereby preclinical drug target identification and validation which occurs in cells, tissues, and animal models of disease are poor predictors of human efficacy. This means definitive evidence on the validity of a new target for the treatment of human disease comes late in drug development, during clinical (phase II or III) randomised controlled trials. Because the drug target hypothesis advanced by preclinical studies is all too frequently false, expensive, late-stage failure due to lack of efficacy is an increasing problem, affecting many therapeutic areas, posing a major threat to the sustainability of the current model of drug development. A seemingly unattainable solution to this problem would be to obtain large-scale randomised human evidence on a target and disease state earlier in a drug development programme, without recourse to developing a medicinal compound that might be destined to fail.

The Research and Results
Our work extended the concept of Mendelian randomisation studies for drug development from individual targets to the whole genome. The study (1) defined a set of genes that not only encode actual (or potential) drug targets but which are also likely to be responsible for genetic associations with complex diseases from prior genome wide association studies (GWAS); (2) led to the design a genotyping array with enriched single nucleotide polymorphism (SNP) coverage of the druggable gene and; (3) linked the proteins encoded by this gene set to licensed drugs or to compounds with bioactivities against these targets. A variety of bioinformatics resources and other in silico tools were used to achieve these aims. The integrity of the analysis was evaluated through a comparison of the consistency between licensed drug indications and GSWAS associations through manual curation and blinded clinical expert review. This analysis showed that GWAS have already ‘rediscovered’ around 70 or so of the approximately 600 targets of licensed drugs through associations with disease indications, disease related biomarkers or mechanism-based adverse effects. The findings suggest the approach has promise as a tool to systematically identify target-human disease indication pairings and to identify drug repositioning opportunities for licensed drugs. Large scale electronic health record resources with and without genetic information are likely to play an important role.

The Impact
The work has highlighted the important part genomics can play in drug target identification and validation, addressing the major reason for late-stage drug development failure. The findings have been influential in directing pharmaceutical industry attention towards the use of genomics for drug target identification (not just drug response prediction), leading to a growing number of pharma-healthcare (particularly those with large scale electronic health records) and pharma-academic genomics partnerships focusing on drug development.

Reference

The Challenge
Atrial fibrillation (AF) is the most common cardiac arrhythmia and anti-coagulant drugs are used to lower the risk of ischaemic stroke in individuals at high risk according to the CHA2DS2-VASc score. How stroke risk and net clinical benefit differ across CHA2DS2-VASc score in women and men has not been evaluated, yet guidelines recommended both women and men with a score of 2 for anticoagulation.

The Research
Researchers from The Farr Institute in London carried out the first large-scale nationally representative study of the potential benefits and harms of warfarin in individuals with AF across stroke risk and across primary and secondary care. The study used primary and secondary care and national registry data in the CALIBER platform (www.ucl.ac.uk/health-informatics/caliber) to identify individuals with a diagnosis of AF in primary or secondary care between 1998 and 2010. Baseline CHA2DS2-VASc scores were calculated by assigning 1 or 2 points for each stroke risk factor defined according to existing CALIBER phenotypes, distinguishing primary care and secondary care sources.

The Results
70,206 individuals with initial record of diagnosis of AF in primary (n=29,568) or secondary care (n=40,638) in England were included in the study. Compared with individuals with initial record of diagnosis in secondary care, those in primary care had lower overall ischaemic stroke incidence (incidence rate (IR) (95% CI)) per 100 person-years: 2.3 (2.2-2.4) vs 4.3 (4.2-4.4), p value=0.00) and lower scores of ischaemic stroke risk (CHA2DS2-VASc≤2: 30.8% vs 20.6%). However among individuals with low stroke risk (CHA2DS2-VASc (0, 1 or 2) there were no differences in ischaemic stroke rate between those with initial record of diagnosis in primary care or secondary care. The overall net clinical benefit (NCB) of warfarin was 1.9 (1.8-2.1) ischaemic strokes avoided per 100 person-years. For CHA2DS2-VASc=0, CHA2DS2-VASc=1 and CHA2DS2-VASc=2, NCB was (NCB (95% CI)/100 person-years: -0.3 (-0.8-0.1)), (NCB (95% CI)/100 person-years: 0.1 (-0.2-0.4)) and (NCB (95% CI)/100 person-years: 0.2 (-0.1-0.6), respectively. Results showed a significant positive NCB of warfarin from CHA2DS2-VASc≤2 in men (NCB (95% CI)/100 PY: 0.5 (0.1-0.9)) and from CHA2DS2-VASc≤3 in women (NCB (95% CI)/100 person-years: 1.5 (1.1-1.9)).

The Impact
Clinical guidelines (European Society of Cardiology 2016) were changed by this research; previously women with a CHA2DS2-VASc score of 2 were recommended warfarin, and based on these sex specific net clinical benefit findings, in the 2016 guidelines this was changed ≤3. Many women will be spared ineffective, harmful anti-coagulation.

Reference

4. Changing Hospital Approaches to Outbreaks of Infectious Disease: Whole genome sequencing of viral pathogens integrated into routine healthcare to support patient stratification and viral outbreak surveillance

The Challenge
Current clinical treatments of viral infections are largely conservative and do not take into account the genetic information of the infecting pathogen. If clinicians could have high resolution information about the gene variations in specific cases of infections, more targeted and effective treatments could be selected. The significant genomic variation of viruses underpins pathogenicity, drug resistance and transmission and allows genetics based patient stratification, infection control and molecular epidemiology.

The Research
ICONIC, a major project funded by the UK Department of Health and the Wellcome Trust, has developed sustainable next generation technology for the viral full-length genome sequencing using residual diagnostic samples within the diagnostic value chain. ICONIC processes and assemblies de novo the virus genomes using high performance computation available in ‘cloud resources’ bypassing the need for embedding complex computer technology in the NHS and specialist staff. It links virus genetics and diagnostic information to hospital based clinical information and generates the appropriate clinical reporting tools.

The Results
ICONIC is embedding this new technology within the UK NHS diagnostic environment. The study team designed methods to process diagnostic clinical samples and a computer-based tool to quickly and accurately reconstruct entire sets of virus genes. During outbreaks very large amounts of genetic information can be processed and made available in time to help clinicians treating individual patients. By incorporating other clinical data, insights from patient outcomes having different treatments can be gained.

The Impact
The novel technology has influenced direct patient management, hospital infection control and epidemiology and surveillance. The ICONIC process has been adopted by two large NHS Trusts, and has been used to investigate influenza outbreaks at a major hospital trust and has been applied in the Ebola and MERS outbreaks in West Africa and Saudi Arabia. Norovirus and influenza outbreaks in hospitals are due to the introduction of a single case but often it is unclear if the outbreak persistence is due to ongoing person-to-person transmission or from community introductions. ICONIC aims to deliver rapid, real-time results on these questions determining optimal infectious control management. Understanding the community spread of influenza and measles infections is essential for implementing appropriate control measures, ranging from local immunisation, to national surveillance.

References
5. Changing Cities: Harnessing the potential for data linkage, a case history from Born in Bradford

The Challenge
Mobilising whole communities in research spanning genetic through to environmental determinants of health, and incorporating frameworks for intervention, has been a challenge.

The Research
The Born in Bradford (BiB) cohort study was established in 2006 to examine how genetic, nutritional, environmental, behavioural and social factors impact on health and development during childhood, and subsequently, adult life in a deprived multi-ethnic population. Between 2007 and 2011 detailed information on socio-economic characteristics, ethnicity and family trees, lifestyle factors, environmental exposures, physical and mental health have been collected on 14,000 women, their children and 4,000 of their partners.

Using NHS numbers with matching to National Pupil Numbers, data linkage has been successfully achieved between maternity, child health, hospital, primary care and education records. This data linkage is now providing rich information to develop health profiles of children in Bradford to provide an assessment of the needs to inform commissioning of services. Paediatricians have been able to use this data to develop a comprehensive register of congenital anomalies. Obstetricians have used the linkage to evaluate the impact of universal screening for diabetes in pregnancy. Health visitors have been able to monitor child growth and obesity and improve targeting of support. Midwives have used the feedback of accurate data to redesign antenatal services. The data linkage has provided a catalyst to develop primary care templates to improve the quality of outcome data.

The Results
BiB has linked research data with routine air pollution monitoring data to demonstrate the harm to foetal growth. Genetic data (exome sequencing) has been linked to primary care data to provide a simple and efficient method of investigating the impact of genetic variants on health.

The Impact
The power of routine data linkage has enabled BiB to establish the world’s first experimental birth cohort study to provide evidence of the effectiveness of 22 early life interventions in some of the most deprived communities in the UK.

Reference


The Challenge
Existing studies evaluating the observed global decline in mortality following non-ST-elevation acute myocardial infarction (NSTEMI) are limited by sufficient sample size and duration of follow-up. The extent to which this decline can be attributed to lower baseline risk among these patients or stricter adherence to guideline-indicated therapies is not known.

The Research
A team of researchers from the UK (Farr HeRC and Farr London), Canada, Spain and Australia used data recorded in the Myocardial Ischaemia National Audit Project (MINAP) registry of all acute myocardial infarction hospitalizations in England and Wales to identify a cohort of patients eligible for the study. Patients aged 18 years or above admitted to hospital with a final diagnosis of NSTEMI between 1 January 2003 and 30 June 2013 were selected for inclusion in the study. Final diagnosis was ascertained by local clinicians according to presenting history, clinical examination and results of inpatient investigations in keeping with the consensus document of the Joint European Society of Cardiology and American College of Cardiology. Clinical factors included baseline Global Registry of Acute Coronary Events (GRACE) risk score, patient demographics, comorbid conditions, medication at discharge, use of an invasive coronary strategy and mortality. Patients who died in hospital or who had missing post-discharge mortality data were excluded from the analysis.

The Results
389,057 patients were included in the study (median age 72.7 years [IQR 61.7-81.2 years]; 63.1% men). The study observed changes in clinical presentation including increased diabetes, hypertension, cerebrovascular disease, use of coronary angiography, percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) and cardiovascular disease medication over the study period. There was a significant decrease in 180-day all-cause mortality from 10.8% to 7.6% that was associated with increased use of an invasive coronary strategy after adjustment for changes in clinical risk and pharmacological therapies.

The Impact
This study is one of the first to describe at scale the high but decreasing baseline acute coronary syndrome risk profile of patients with NSTEMI as well as their increasingly comorbid status. Improvements in all-cause mortality were observed over the 10-year study period up to 2013. Increased mortality was associated with use of an invasive coronary strategy and not entirely related to a decline in baseline clinical risk or increased use of pharmacological therapies

Reference
7. Changing a Country: Wales as a national public health laboratory-improving mental health through the regeneration of deprived neighbourhoods

The Challenge
Neighbourhood-level interventions provide an opportunity to better understand the impact of neighbourhoods on health. In 2001, the Welsh Government funded Communities First, a programme of neighbourhood regeneration delivered to the 100 most deprived of the 881 electoral wards in Wales. In this study, the authors examined the association between neighbourhood regeneration and mental health.

The Research
Information on regeneration activities in 35 intervention areas (n=4,197 subjects) and 75 control areas (n=6,695 subjects) were linked to data on mental health from a cohort study with assessments in 2001 (before regeneration) and 2008 (after regeneration). Propensity score matching was used to estimate the change in mental health in intervention versus control neighbourhoods. Baseline differences between intervention and control areas were of a similar magnitude as produced by paired randomization of neighbourhoods.

The Results
Regeneration was associated with an improvement in the mental health of residents in intervention areas compared to control neighbourhoods, suggesting a reduction in socioeconomic inequalities in mental health. There was a dose response relationship between length of residence in regeneration neighbourhoods and improvements in mental health.

These results show the targeted regeneration of deprived neighbourhoods can improve mental health. Also, the Communities First regeneration programme evaluated in this study is unique in that community residents, rather than local councils or governments, identified areas to be regenerated. Targeted regeneration directed by the residents of deprived urban communities, may therefore help to reduce inequalities in mental health.

The Impact
The study suggests that targeted regeneration of deprived neighbourhoods can improve mental health.

Reference

8. Citizen Driven Health: Recognising social functioning from smartphone GPS data in patients with serious mental illness

The Challenge
Risk of psychotic relapse in patients with schizophrenia is commonly measured through social functioning, which focuses on patients' daily activities. Monitoring of social functioning usually relies on infrequent clinic visits, limiting the capacity to detect sudden changes. Global Positioning System (GPS) data that is passively collected with smartphones introduce new opportunities to timely monitor aspects of social functioning related to out-of-home activities. A pilot-study was conducted by researchers at Farr HeRC to assess the feasibility of this approach.

The Research
For five days, five patients with schizophrenia continuously recorded their GPS location with a smartphone and completed a paper-based social functioning diary to register their out-of-home activities.

A time-based method and a density-based method were implemented to identify the geolocations visited from the raw GPS data. Then the geolocations visited were clustered into places visited. Finally, we used semantic enrichment to identify each participant’s home and work place and to classify the remaining places and associated activities. For example, places like cinema and theatre were associated with "recreational activities", while pubs or restaurants were linked to "social activities". Accuracy was assessed by comparing the activities detected from the GPS data with those recorded in the social functioning diary.

Recall ranged from 0.657 (Standard Deviation [SD] 0.190) to 0.771 (SD 0.264) while precision ranged from 0.661 (SD 0.197) to 0.954 (SD 0.093). As expected, recall improved when the time threshold (i.e. the time GPS data needed to be within a certain radius to be considered a geolocation visited) decreased, while precision increased when the time threshold increased. For the same time threshold, the density-based pipeline presented better recall, while the time-based method obtained better precision.

The Impact
This study shows how using GPS data and relatively simple analytical methods can be used to detect patients’ out-of-home activities with moderate recall and precision. It demonstrates the potential of using routinely and passively collected GPS data from smartphones to assess out-of-home behaviours, which are crucial to monitor patients with mental health issues.

Reference
Enabling Research at Scale: datasets, infrastructure, methods, standards. governance

Objective 2; Enable new datasets and develop new infrastructure, methods, technologies and standards for such research.

Objective 4; Work with the owners and controllers of data to support the safe use of patient information for medical research across the UK, championing data protection, confidentiality and privacy.

Trusted Research Environments

The Farr Institute capital investments for physical safe haven nodes in 2015 progressed to digital operations in 2016 on common standards that will support collaborative analytics in the future.

Joining up the UK

The Jisc Safe share project is working with The Farr Institute, the MRC Medical Bioinformatics initiative, and the Administrative Data Research Network (ADRN), piloting a high capacity, encrypted network between research centres across the Janet network. Jisc is now working with local authorities and the Cabinet Office to extend the reach of the Public Services Network (PSN). The Higher Assurance Network (HAN) equipment which has been installed in Farr CIPHER and HeRC was successfully tested and is now in routine use at Farr CIPHER.

The Safe share Authentication Authorisation and Accounting Infrastructure (AAAI) component will be further developed through a new National AAAI Pathfinder project. This project, funded through contributions from all seven Research Councils, will create a common user authentication and management platform for compute and data resources. Partners include Edinburgh Parallel Computing Centre, University of Leeds, The Francis Crick Institute, University of Oxford, Science and Technology Facilities Council and is led from University College London.

The Farr Institute and The Alan Turing Institute held a joint workshop in July 2016 to explore common interests in accessing and analysing data in secure environments. Over 20 attendees from eight institutions attended as well as representatives from Intel’s Health and Life Sciences team. The Farr Institute and The Alan Turing Institute have both agreed to jointly fund a new project, FASHIN: Federated Access to Secure Healthcare Information (Pis: Dr Efi Tsamoura Turing Research Fellow, Prof Dave Robertson, Farr Scotland,) to develop techniques that will provide more automation in running queries over The Farr Institute’s infrastructure, reducing the amount of computational resources required to run experiments. The algorithms that will be developed will be implemented in a prototype and demonstrated using data from NHS and query workloads from The Farr Institute.

National and Regional Trusted Research Environments

In Scotland there is joint national infrastructure that has been funded over the last two years by The Farr Institute, the Administrative Data Research Centre in Scotland, and the Scottish Government’s Chief Statistician’s office. This year the Urban Big Data Centre (funded by ESRC) has joined the collaboration and has begun contributing to the shared infrastructure. This model of pooling resources and processes has been very effective and will allow the service to run in its current form until the end of 2018. Collaboration is at the heart of Farr Scotland with the data linkage service being provided by National Records Scotland, the hosting and management of the data analysis by the Edinburgh Parallel Computing Centre and research support and project management by eDRIS, which is part of National Services Scotland within the NHS. Scotland also has four regional safe havens: two have received accreditation from NHS Scotland eHealth, one is in the process of applying for accreditation and one has decided to use the National Safe Haven in the short to medium term. Two also have ISO27001 Information Security accreditation. All regional safe havens continue to support local research activity. At national level Farr Scotland has over 250 live projects, with around 70 of these at the analytic stage involving over 120 academic staff. The profile of Scottish projects shows that 75% are academic led, 20% NHS or Scottish Government, 3% commercial, and 2% third sector.
The Farr HeRC Trusted Research Environment (TRE) is now integrated into the Greater Manchester Datawell, an integrated electronic health record for primary, secondary and tertiary care, covering a 2.8 million population. The Farr Institute HeRC TRE provides the data and analytics infrastructure for the Greater Manchester Connected Health City and plays a key part in data analytics for Greater Manchester health and social care devolution. The Farr HeRC TRE is now an active component of the health system it is connected to, acting as a hub through which data can be exchanged between NHS sites, research infrastructure and patients. Across the North of England, work is ongoing to network the data analytics centres of the four Connected Health Cities with the Farr Institute HeRC TRE, creating a federation of population health centres across the North of England.

The three London Academic Health Science Centres have come together to develop a London Trusted Research Environment. Funding has been obtained to recruit a dedicated project manager to support the project. Farr Institute representation on the Delivery Board includes Prof Andrew Morris (Chair, UCLP Informatics Board) and Prof Harry Hemingway. Dr Daniel Ray (Director of Data Science, NHS Digital, and Farr London Honorary appointment) is involved with the project ensuring that there are strong links with NHS Digital.

The UK Secure e-Research Platform (UKSeRP) hosted by Farr CIPHER is now fully operational and currently provides ISO27001 certified data curation, management, sharing, and analysis facilities to the UK research community. Imbued with capabilities derived from 10 years of operating the Secure Information Linkage Databank (SAIL Databank) in Wales, UKSeRP is provisioned with a suite of capabilities that allow data owners to remotely share their data in line with best information governance principles using an easy-to-use interface in a ready to use solution. UKSeRP supports the Dementia Platform UK (see case study page 10), the SAIL database, the platform’s biggest user, UK Biobank, the UK Multiple Sclerosis Register, the Administrative Data Research Centre Wales (ADRCW) and the MRC Medical Bioinformatics CLIMB project (Cloud Infrastructure for Microbial Bioinformatics). Requirements of ADRCW have led to an instance of UKSeRP with enhanced security controls to meet the expectations of UK Government departments. Closely coupled to a number of “safe rooms”, data is stored and accessed via a CLAS (CESG Listed Advisor Scheme) consultant approved technology set, with capabilities that are now available to any UKSeRP tenant wishing these enhanced information architecture and security features and accreditations.

The Dementia Platform UK (DPUK) has selected UKSeRP as the unifying Data Portal to bring together and share data from 33 pre-existing research cohort studies.

With robust legal agreements for data transfer now in place with five academic institutions and an industry partner, data is now being received in Swansea from cohorts across the UK. Three cohorts have currently shared baseline study data to the Portal. Another six are currently online and using our front-end to the infrastructure to upload their data and supporting documentation or are in advanced discussions for sharing multi-omics data. After a short period of testing, this data will be made available to the wider research community, subject to successful project applications. It is envisaged that 16 cohorts which are currently deemed ‘research-ready’ will have shared data by June 2017 and preliminary feasibility and case study projects are due to start in order to provide operational evidence of the capability and functionality of the DPUK Data Portal.

New UKSeRP infrastructure builds include the integration of the DPUK genomics platform to allow for centralised genomic research to take place within a specialist analysis platform on UKSeRP, along with epidemiological, imaging and wearables data.

In partnership with the European Medical Information Framework (EMIF), DPUK will be the recipient of extra metadata tools for in-depth cohort discovery, and subsequently, an integrated participant finder tool that will be enhance the capability of DPUK cohorts to be used to recruit for clinical trials, which is of particular interest to the European Prevention of Alzheimer’s Dementia (EPAD) consortium. UKSeRP has recently added XNAT, an open source imaging informatics platform, to support DPUK’s needs, as well as making it available to other tenants.
Enabling Research at Scale: datasets, infrastructure, methods, standards, governance

**Discovery Across Data Sectors**

In this reporting period The Farr Institute has delivered data-discovery facilities across different regional, administrative or discipline-based datasets.

Following The US-UK Transatlantic Health Data Science summit in March 2016 (sites.google.com/site/usukhealthdata/home), there has been increased collaboration in this field, including the award of a $100 million NIH grant to bring 17 birth cohorts into a single digital laboratory that can take new Fast Healthcare Interoperability Resources (FHIR) compliant data feeds, as well as support collaborative analytics with rich, curated metadata. The ‘Children’s Respiratory Research and Environment Workgroup’ (CREW) funded through the National Institute of Health is to be developed on the Farr HeRC (STELAR) eLab infrastructure to allow research data to be combined across the US and UK (initially 17 studies) creating, for the first time, an asthma dataset that is rich enough to allow the complex analysis required to significantly improve our understanding of asthma and underlying mechanisms. Importantly, the collection of data from FHIR compliant feeds opens up an opportunity to support direct data collection from mobile health devices.

Farr Institute engineering has applied the concept of Research Objects underpinning the eLab to support the FAIR (Findable, Accessible, Interoperable and Reusable) use of data through collaboration with the Centre for Longitudinal Studies (CLOSER), UCL.

Farr London have established an on-going collaboration with the Federal University of Bahia, Brazil (115 million cohort) and Prof Mauricio Barreto is visiting Farr London through a Royal Society Newton International Fellowship to develop novel analytical methods for linkage of large administrative and health record sources.

Involvement with the MRC-funded Cloud Infrastructure for Microbial Bioinformatics (CLIMB) has allowed UKSeRP to be augmented with significant High Performance Computing capabilities to support genomics analyses. Operating fully as part of the UK-wide CLIMB consortia, the technology investment has been accommodated into the existing UKSeRP infrastructure to also allow unused cycles to be repurposed for other studies and tenants, as well as providing the necessary secure circumstances to allow the co-analysis of microbial and human genetic data, along with clinical and population data.

Farr London members at QMUL have established a robust Informatics Hub Infrastructure to support a range of MRC stratified medicine projects (MRC-RA-MAP (RA disease biomarkers), MRC-MATURA (RA Biologic response), MRC-PSORT (Psoriasis Biologic response), MRC-TARGET (Giant Cell Arteritis biologic response). Complex multi-omic data on several thousand Rheumatoid Arthritis (RA) and Psoriasis patients has been captured and integrated within our tranSMART infrastructure hosted on MRC eMedLab, the Medical Bioinformatics private cloud infrastructure. This data formed the basis of an MRC-sponsored pan-Immunology disease working group (13th Dec 2016) to plan opportunities for data integration across the stratified medicine. As an outcome of the meeting, the QMUL-Farr team are collating data fields across all Immune stratified medicine cohorts, in collaboration with the MRC Leicester pathology node, the Leeds Medical Bioinformatics Centre and Farr HeRC.

**Analytic Methodology**

The Farr Institute has continued to push analytic boundaries over well-formed problems and to bring together biostatisticians, mathematicians and computer scientists into an open-minded health data science colloquium, extending to MRC/EPsRC Molecular Pathology, MRC Stratified Medicine and MRC Medical Bioinformatics awards.

At the interface of machine learning, statistics and epidemiology, The Farr Institute has taken a deep dive on the challenge of utilising rich yet irregular longitudinal data for inference. One specific output has been the uncovering of latent trajectories of change over time in Body Mass Index (BMI), which has identified subgroups of the population at higher risk of some cancers (e.g. press coverage www.dailymail.co.uk/health/article-3911474/How-middle-age-spread-increases-cancer-risk-50-Men-need-gain-2st-7lbs-teenage-years-retirement-raise-risk.html).

Funding from the ESRC and Welsh Government has provided resources to roll out The Farr Institute’s National Research Data Appliances (NRDA) into nearly half of the local authorities of Wales. The NRDA have recently augmented the on-board natural language processing facilities to address the significant barriers to efficiently finding eligible participants for clinical trials. Piloted in a local Welsh health board, the technology set consumes all of the documents (letters, reports and freetext database entries) generated by any part of the hospital, converting the content into fine-grained SNOMED CT terms. An intuitive user interface then allows hospital staff to articulate clinical trial inclusion and exclusion criteria and run the resultant query against the data. Pilot work has shown this increases the number of patients found, and decreases the time taken to identify them, creating massive advantages for any study site. This pilot work is currently the subject of a proposal to Welsh Government to implement this solution across the whole of Wales’ hospitals, providing a comprehensive set of clinically-rich, all-Wales data to the SAIL Databank.

CALIBER, a Farr London research resource of linked electronic health record data in 10m patients from primary care, hospital care and mortality records was selected as one of three exemplar UK research projects showcased at the National Institutes of Health in Bethesda at the NIH/RCUK Transatlantic Data Science Workshop. This led to the establishment of new transatlantic methodological collaborations (Prof Hernan, Harvard) and multiple NIH grants submitted.

**New Data Feeds**

The Farr Institute has prepared its trusted research environments to receive new data feeds, for example from multi-omics, mobile sensors and ‘digital self’ ubiquitous data aggregations.

The Farr Scotland data analysis platform has continued to evolve over the past 12 months, moving from the simple single secure zone approach which went live in December 2015 to a multi-zoned approach that separates the active research projects from data such as the PACS (picture archiving and communication system) imaging database. Over the past 12 months, a link into NHS Scotland’s live PACS systems has been established which will allow, over the next 18 months, a full copy of the past 6 years of medical imaging data to be retrieved and pseudonymised for subsequent scientific analysis. This will be a unique resource once the data retrieval process has run its course. The team from the
Enabling Research at Scale: datasets, infrastructure, methods, standards, governance

Health Informatics Centre at the University of Dundee have continued to develop the image management software which will be used to manage the image repository and provide safe researcher access to it.

There is a clear need to unlock the deeper data within the unstructured component of the electronic health record. To address this, in collaboration with King’s College London (Prof Richard Dobson, joint appointment KCL/UCL Farr London), we have developed CogStack, a toolkit that unlocks the unstructured information for real time decision support, tailored care and recruitment into trials and Genomics England’s 100,000 Genomes Project, ‘Google-ising’ the record for care, BI, audit, coding and through optional de-identification, research. This is transformative in the Trust sites where it has been deployed to date with semantic annotation and search capability overlaid, based on data standards such as SNOMED and Human Phenotype Ontology (HPO), and delivered through the Horizon 2020 KConnect project (kconnect.eu). Other unstructured text based research programmes include 100,000 radiology images and reports used to build Deep Learning based models of outcome prediction in dementia. Additional projects include the D-CRIS programme, a de-identified version of the Camden and Islington mental Health electronic health record. The Farr Institute is working with The Alan Turing Institute and Intel, with Turing Fellows Dr Mirco Musolesi (UCL) and Dr Maria Wolters (Edinburgh), to build on the Cogstack platform (£22 million IMI-funded RADAR-CNS).

The ‘Wearable Clinic’ is an EPSRC-funded collaboration between Farr HeRC partners that deploys emerging digital technologies to help patients with long-term conditions to better manage their health in daily life, respond more quickly to changes in symptoms and prevent fall back episodes. The project will trial methods for (1) acquisition of real-time, high resolution data from wearable sensors; (2) dynamic prediction of care needs and health outcomes; (3) creation of highly-personalised care plans that adapt to predicted needs/outcomes; and (4) understanding how this will impact on patients, their carers, care professionals, and care services to inform strategies for adoption. The project will build on two clinical exemplars; serious mental illness (in particular, schizophrenia) and chronic kidney disease. Project partners are Cerner Corporation; Withings SAS; UK Renal Registry; Manchester Mental Health & Social Care; the National Institute for Health and Care Excellence (NICE); NHS Digital; Manchester Connected Health Ecosystem; and Health Innovation Manchester.

SAIL Databank is increasing its holdings in health determinants and outcomes data with sharing agreements in place with organisations such as South Wales Police, the National Offenders Management Service, local government social services and housing departments and many more. Farr London investigators, in collaboration with the ADRC–England, have developed methods for mother-child linkage of 10 million deliveries in England and are working with Public Health England and the Ministry of Justice to apply new linkage methods to government-held data.

Data Resource Profile: The Scottish National Prescribing Information System (PIS)

Challenge. To describe for the first time to researchers worldwide the data available through the new Prescribing Information System (PIS) on prescriptions prescribed, dispensed and reimbursed nationally in Scotland within the community. To indicate how data are generated, how it can contribute to international research studies, its strengths and weaknesses and how it can be accessed.

Methodology. The Farr Scotland Pharmacoepidemiology Group developed a specification defining researchers’ core data needs which informed a development programme for enhanced provisioning of PIS, delivered in collaboration with Information Services Division, NHS National Services Scotland.

Results. PIS includes information about patients, prescribers and dispensers’ characteristics and the drugs prescribed, dispensed and reimbursed. Since 2009 the records in PIS are available at an individual level (over 500 million items prescribed from 2009 to 2014) and since 1993 at an aggregate level for the 5.3 million residents in Scotland enabling linkage to local and national databases.

The data is derived from three sources: the ePrescribed messages (generated by General Practitioners), eDispensed messages (generated by community pharmacies) and Reimbursed messages from scanned paper prescriptions dispensed in pharmacies. Natural language processing methods have been applied to transform drug dosage instructions into structured fields to support quantification of drug exposure over time. Examination of medicine adherence through suitable standard tool application for these data have been progressed.

Conclusions and Impact. This new dataset is available upon request through the electronic Data Research and Innovation Service ([NSS.eDRIS@nhs.net]) and is already being used to support studies on the utilisation of drugs (new and established therapies including antimicrobials, cardiovascular, diabetes, methadone), measurement of polypharmacy improvement programmes, and supporting generation of risk assessment/predictive analytics tools to support clinical decision making.

Next steps: This work has supported the award of a European Medicines Agency preferred partner status for pharmacovigilance studies and is being positioned to enable Scotland to be a global partner in new real world evidence generation on the safe and effective use of medicines.

**Actionable Analytics**

The Farr Institute has developed further exports of digital entities that feed NHS-based analytics with methodology and evidence from university-based trusted research environments.

The infrastructure group are engineering a bidirectional discovery-implementation environment for a future of increasing frequency of observation where discovery is provoked by designed and natural experiments in real-world healthcare. The capability to write back into electronic health records has been demonstrated through the ARUK REMORA (Remote MOntoring in Rheumatoid Arthritis) project (www.clahrc-gm.nihr.ac.uk/our-work/exploiting-technologies/remora/). Innovative forms of data have been supported through the NIHR Cystic Fibrosis Health Hub, enabling adherence data to be captured directly from inhalers and fed back to patients and healthcare professionals to enable coproduced health, linking into the NIHR Cystic Fibrosis health observatory. In the spirit of innovation, a prototype of a citizen’s portal which allows patients to set preferences on how their data is used, has been built on blockchain technology, enabling a full record of data transactions to be kept. This prototype has been adopted by Connected Health. Cities as the foundation for a citizen’s portal, enabling control and feedback on data usage.

The NIHR Health Informatics Collaborative (NIHR HIC). The UCL-led Critical Care theme has established itself as the leading theme from the five scientific themes initiated, resulting in a community of researchers sharing and developing capacity through ‘datathons’ (www.datascicc.org) exploiting the ability to share the data beyond the UCL Data Safe Haven through k-anonymisation. Data processing pipelines, data cleaning and data analysing tools are available as open source components (https://github.com/CC-HIC/ccdata).

The Farr CIPHER backed Prudent Healthcare Intelligence Hub, based at Swansea University but working across the whole of the Welsh NHS, delivers analytical services to support the direct needs of healthcare organisations and policymakers by undertaking high impact analyses and evaluations. With involvement of staff from across the Welsh NHS, this cooperative model, underpinned by Farr Institute infrastructure and SAIL Databank data, is proving a highly efficient means of gaining real impacts into society, as well as building capacity and capability in the NHS.

Farr Institute researchers are working closely with NHS England, and NHS Digital, on a Digital Interoperability Target Architecture, drawing on the place-based learning health systems work that underpins the Department of Health Connected Health Cities pilots. A core principle of ‘diameter of trust’ has been adopted, whereby regions are large enough to achieve economies of scale in digital infrastructure yet small enough to involve their communities sufficiently to earn a civic mandate for deeper analytics than the public would trust to external parties. The optimum size of such regions in England is estimated at 1-5 million population. This work has considered not only the interoperability of digital systems within regions but also how regional and national systems operate in synergy. NHS operational needs are very similar to research needs for data and analytics when considered in the framework of learning health systems, and this joint thinking has led to considerable recent progress toward an efficient regional digital health federation for England.
Capacity Building

Objective 3; Develop skills, talent and expertise in individuals and research communities for collaborative working.

Capacity building continues to be a key focus of The Farr Institute. Health informatics is exceptional in the number of disciplines required to contribute to the successful delivery of impactful research. This brings both advantages and challenges to creating training and education programmes which allow individuals to develop their careers within this field. There is not a single trajectory or career path along which professional development can be mapped. The Capacity Building Work Group (CBWG) have embraced this challenge and designed a strategy which allows for individuals with different experiences and backgrounds to learn together in a community of practice in which individuals can derive new knowledge and skills from both instructors and peers.

The Farr Institute Doctoral Training Programme

The Farr Institute’s Centres have been facilitative in creating a positive environment for attracting further funding to create PhD studentships. A number of The Farr Institute’s member universities have been successful in obtaining Doctoral Training programme status from RCUK funders [e.g. MRC Doctoral Training Programme in Precision Medicine (Scotland) and the ESRC/BBSRC Centres for Doctoral Training in ‘Biosocial research’ (HeRC and London) and ‘New forms of data’ (HeRC)]. There are now over 100 PhD students working across the many disciplines contributing to health informatics research aligned with the four Farr Institute Centres. The CBWG continue to organise symposiums and summer schools to develop a peer-led community of practice at a UK-level.

The second Farr Institute PhD symposium took place at Farr London 23 - 24 May 2016. Twenty-one students representing the multi-disciplinary nature of Farr Institute research activities came together from across the UK for the opportunity to network, exchange ideas and apply their knowledge to solving real-world challenges.

The theme of this year’s symposium was ‘sandpit’ methodology. The aim was to bring cross-discipline expertise together in a free-thinking environment in order to co-develop solutions to a problem. The students were provided with real-world data challenges from AstraZeneca, Patients Know Best, Public Health England, NHS Choices and Cerner. The case studies posed problems each organisation was currently facing including:

- dynamic patient consent
- improving analysis of patient feedback to improve the health service and the patient experience
- modelling the health benefits of helping people continue to work in their later years
- utilising the electronic health record to identify the actual use of approved oncology drugs to reveal effectiveness and gaps within the treatment pathway

The Farr Institute PhD Symposium, London 2016.
Capacity Building

- application of digital systems to enhance and encourage citizen engagement in their healthcare
- Reflecting on his experience of the event, Glen Martin, PhD student at Farr HeRC said “working with healthcare and industry leaders around interesting and difficult real-world problems was extremely beneficial. It gave us an insight into how our skills as a team could be applied in the real-world setting”.

An International Summer School was held in conjunction with the International Population Data Linkage Network (IPDLN) Conference 2016 widening the networking opportunities and knowledge exchange available to The Farr PhD students by involving international counterparts. The two-day summer school hosted over 30 students who were introduced to two subjects.

- Day 1: Incorporating open data and unstructured data from apps and wearables in your research
- Day 2: Communicating your research to the public through the media.

The media workshop was provided by freelance journalist Ms Pennie Taylor who has worked for BBC Scotland and The Herald newspaper. Ms Taylor provided a fascinating and enlightening perspective on the aims of the media in reporting on medical research and how to work with this to try to ensure the conclusions of research are accurately represented. The students had the opportunity to practice writing their thesis research as a 500 word article for a newspaper and presented this back to the wider group.

Continuing Professional Development

The Farr Institute recognises the need to develop capacity through provision of training and education to those already in the workforce across academia, industry and healthcare providers. Farr Institute partners have continued to invest in and develop short courses and Masters qualifications which can be taken on a part-time basis. In 2016 Farr HeRC ran the first Farr MOOC (Massive Open Online Course) aimed at clinical bioinformaticians. The MOOC had two runs in June and September 2016. The course lasted 5 weeks and the June run enrolled 5,150 students with 1,745 active learners. Those enrolling were a mixture of students, professionals in the field and general public with an interest in the area.

Each Farr Institute Centre continues to offer short courses (1 day - 1 week) to allow participants to increase their skills and knowledge in particular areas. Farr London continues to run its 2 week short course programme covering all aspects of electronic health record research, giving the 133 participants in 2016 the opportunity to mix and match courses to suit their personal needs.

The Farr HeRC education team continue to work closely with NHS colleagues in the North West Strategic Health Authority to train NHS staff in the informatics skills to manage the data accumulating within the NHS service. In addition to the ‘Informatics for Health Care Systems’ programme one day workshops on statistical methods and data modelling trained 58 NHS staff in 2016/17.

Farr CIPHER have formatted their MSc in Health Informatics so that it can be taken as a full or part-time MSc or students can choose to study a particular module as CPD. NHS and industry internships form part of the dissertation and work-based project of this course. In partnership with the National Health Service (NHS) Wales Informatics Services and Health Boards in Wales, this course is able to offer NHS research opportunities within the NHS Wales Informatics Research Laboratories based at Swansea University. Over 100 students have now registered for the MSc/MRes and 238 professionals have completed individual modules as part of their continuous professional development.

A full list of the Masters courses offered across The Farr Institute can be found on our website: www.farrinstitute.org/research-education/education/post-graduate-programmes

NHS Collaborations: Modernising Scientific Careers & NHS Digital Academy

Modernising Scientific Agenda around Health Informatics continues to be delivered by University of Manchester and UCL. The Higher Specialist Scientific Training curriculum contributed to by Farr Institute academics and led by George Moulton has been finalised and will be going live in 2017/18. In addition, work is continuing to address lower career levels of clinical scientists in the broader area of clinical bioinformatics also contributed to by The Farr Institute at the University of Manchester.

Since January 2017, The Farr Institute has contributed to the scoping of the NHS Digital Academy (announced in September 2016 by Jeremy Hunt in response to the Wachter Review and two of the 10 recommendations that focussed on the development of a clinically engaged IT workforce to engender digital change in the healthcare sector). Farr Institute partners sit on the steering group that have guided the development of the Digital Academy set up and curriculum, and a number of Farr Institute associated universities are preparing joint bids to work with NHS England as the delivery partner for the academy (bid winners to be announced May 2017). Bids will focus on a cross-sector approach to drive digital transformation across the NHS by developing current and aspiring Clinical Information Officers and Chief Clinical Information Officers.

The Farr Institute Future Leaders in Health Data Science

The aim of this programme is to create a cohort-based mentorship and networking programme for researchers who have already demonstrated innovation in the field of health data science to provide them with the cross-discipline research skills, international networks and leadership skills they will require as future leaders in their field.

The programme runs over 2 years with a cohort of 12 successful applicants meeting in each year over a week long residential course. Activities are designed to promote group work and collaboration with the aim of everyone undertaking exchange visits over the 2 year period. The membership of the cohort is carefully chosen to ensure a range of disciplines, expertise and research interests are included to facilitate members identifying fellow cohort members with whom to develop new research collaborations based on their complimentary skills.

The programme was launched at The IPDLN Conference, August 2016, and received 63 applications. A breakdown of the application demographic is shown in Figure 2. The applicants were judged by a panel consisting of the CBWG leads and the Farr Institute Directors. The following markers of esteem were taken into consideration: cited peer reviewed...
publications; contribution of new methods, tools, analytics; input into policy; Principle Investigator or Co-PI on grant awards; Individual Fellowships; demonstration of public involvement and engagement; contribution to national and international networks and organisations. Each applicant submits a personal statement, CV and a letter of support and recommendation from their Head of Institute.

An International programme: It is critical to the success of the scheme that the members of the cohort represent the International research base in this field. Health data science is a relatively recent research discipline and there is not yet a critical mass of people within one country to expose cohort members to the new ideas, experiences and best-practice across this multi-disciplinary field which will help them drive innovation and widen the impact of research outputs that can have direct effects on improving the health of populations. The experience and learning of the UK-based residents on the scheme will be enriched by meeting with their peers from other countries and will bring new learning into the UK, which can speed up innovation and application of best practice within the UK. In addition, it is essential to the career development of researchers that they are able to demonstrate strong international research collaborations and the impact of their research at both a national and international scale.

Financial model: The Farr Institute is funding the cost of the residential meetings at an estimated cost of £2,200 per individual across the 2 year programme. The successful candidates have to commit to funding travel to the residential meetings and costs associated with the exchange visits between cohort members. It is hoped that if the programme is able to build a strong reputation over the next few years, the Institute would then be in a position to charge fees for the programme (similar to the model used to establish the Oxford International Primary Care Research Leadership programme).

Future Plans: Early Career Researcher Exchange Programme

The Memorandum of Understanding signed between The Farr Institute and the Institute of Clinical Evaluative Sciences, Ontario (ICES) (see page 24) provides an opportunity to create a research collaboration based exchange programme. This would be aimed at facilitating early career researchers from both The Farr Institute and ICES to visit each other promoting the transfer of knowledge, skills and best practice between the two organisations. The CBWG plan to work with The Farr Institute Directors and Research theme leads to identify two or three strategic research projects which could be strengthened through a collaboration between ICES and The Farr Institute. This would then provide the basis for selecting early career researchers who would mutually benefit through short-term (weeks/months) research exchange visits. Funding would be provided by The Farr Institute’s Network budget to cover travel and subsistence costs during the visit.

Figure 2. Breakdown of origin of applications for The Farr Institute Future Leaders in Health Data Science
Objective 4: Work with the owners and controllers of data to support the safe use of patient information for medical re-search across the UK, championing data protection, confidentiality and privacy.

The Innovative Governance Working Group (IGWG) has again had a successful, productive year in influencing the health data governance agenda to support the optimum and safe use of data. Group members have been active in contributing to the literature and debate, acting as The Farr Institute’s voice in national consultations, engaging with stakeholders and the public, and working collaboratively towards practical solutions for acceptable data sharing.

Influence

The group responded to the National Data Guardian consultation with detailed feedback that provided commentary on the legal and ethical ramifications of the review, and a discussion that offered points about understanding information security practicalities and risk management (led by G Laurie). Members provided input into Better Data in Government (G Laurie in collaboration with the ADRN), and to the Royal Society and British Academy review of health data governance (K Jones).

Advice on data governance models has been provided to the Irish Health Research Board (HRB) on the development of their national data sharing model (DASSL). It is encouraging that the innovative governance work of the Institute (SAIL and SHIP) was strongly endorsed in the report produced by the HRB.

Members engage with data owners and regulators to work towards acceptable governance solutions for the safe use of patient information for research and contribute to various workshops and symposia on governance issues around the safe use of patient data for research. This includes advising on governance issues pertaining to the use of emerging data types and new infrastructure models (All), and collaboration with the Wellcome Trust on a project examining data custodians’ attitudes towards ‘ownership’ in data sharing (G Laurie).

There have been numerous invited presentations as data governance experts. Some examples are: the EuroCASE annual conference (N Lea and K Jones, Nov 2016), The University of Essex TalkBigData Symposium “Big Data, big brother?” (N Lea), BioData World Congress (K Jones, October 2016), an Information Governance Summit hosted by the Scottish Chief Medical Officer to lay out next steps in information governance for Scotland (G Laurie, October 2016) and proposing a SHIP/Farr Institute principles based approach at an ESRC seminar series (G Laurie).

The international case study (led by K Jones) entitled “The other side of the coin: harm due to the non-use of health-related data” is being used as a background paper to inform the House of Commons Science and Technology Committee’s inquiry into research integrity (February 2017).

Tools and Solutions to Enhance Data Sharing and Use

The IGWG group is strongly instrumental in guiding on tools and solutions to enhance safe data sharing and use.

Working with colleagues at ADRN, a decision-making matrix has been developed to help organisations and institutions assess their own readiness for exploiting existing data and enabling new datasets. Video: www.youtube.com/watch?v=9H1wpoAwmZc and weblink: www.masoninstitute.org/our-research/administrative-data-research-centre-scotland-adrc-s/ (G Laurie)

The group is working in collaboration with the ADRN on a data retention/reuse task group and have produced a principles document on how to proceed on this issue, to build on Farr Institute principles (G Laurie and K Jones).

The Farr Institute Network funded ACoRN project (Advancing Cross-centre Research Networks using anonymised person-based data in data safe havens) began in January 2017. The aim of ACoRN is to scope the barriers to safe data sharing across Centres, to identify practical solutions and, importantly, to explore the drivers that promote crosscentre research. ACoRN includes the work on data sharing between SAIL and eDRIS. The group will be running a workshop at the Informatics for Health conference in April 2017. The session invites an international group of multidisciplinary participants to come together for an opportunity to contribute their expertise to the discussion on the barriers, pitfalls, drivers and examples of good practice in crosscentre research. A report to The Farr Institute Directors will be produced in due course (led by K Jones).

Work has been completed to create a Quality Seal of Approval for Clinical Research in partnership with European Institutes EuroRec and i–HD, and European Medical Information Framework (EMIF). This entails an independent assessment of governance and security best practices for data platforms and organisations that handle clinical data for research purposes. Work is continuing on the development of the anonymisation code of practice in collaboration with NHS Digital and the Information Governance Alliance and on the
development and refinement of codes of practice for reusing clinical data in research across Europe (N Lea).

Engagement

Engagement with all stakeholders is considered to be key to good data governance and the group has been involved in a variety of activities. Some examples are given here:

- Involvement in a review of record linkage approaches across UK and European partners in the EMIF consortium. This has involved interviewing data curators across EU sites about their approaches to linkage and the surrounding legal and technical frameworks: EMIF IMI deliverable completed. A public report due is in March 2017 (N Lea).

- In collaboration with the ADRN, a YouTube video has been produced on the challenges and benefits of data linkage for the public good, focussing on administrative data: https://www.youtube.com/watch?v=9H1wpoAvmZc (G Laurie)

- The group took part in a joint The Farr Institute / ADRN Symposium at the International Association of Bioethics World Congress (G Laurie and N Lea, June 2016)

- Members provided advice to Jo Churchill MP on the National Data Guardian Bill (N Lea and K Jones)

- The group participated in the British Science Festival explaining Big Data and data governance to the general public (K Jones, Sept 2016); a London school science fair that described phenotyping and genotyping (N Lea); and a Tenovus-led public debate on genomic data privacy (K Jones, June 2016)

Dissemination

A key objective of the IGWG is ensuring that the outcome of research is widely shared and that it adds to the general body of global thought leadership on innovative governance is. The following are key papers published by the group in the last year.


- Stevens LA, Dobbs C, Jones KH and Laurie G. ‘Dangers from Within? Looking Inwards at the Role of Maladministration as the Leading Cause of Health Data Breaches in the UK’ in Ronald Leenes and others (eds), Data Protection and Privacy: (In)visibilities and Infrastructures Springer International Publishing 2017


---

Innovative Governance

- Stevens LA, Dobbs C, Jones KH and Laurie G. ‘Dangers from Within? Looking Inwards at the Role of Maladministration as the Leading Cause of Health Data Breaches in the UK’ in Ronald Leenes and others (eds), Data Protection and Privacy: (In)visibilities and Infrastructures Springer International Publishing 2017


---

National Data Guardian for Health and Care

Review of Data Security, Consent and Opt-Outs

- Stevens LA, Dobbs C, Jones KH and Laurie G. ‘Dangers from Within? Looking Inwards at the Role of Maladministration as the Leading Cause of Health Data Breaches in the UK’ in Ronald Leenes and others (eds), Data Protection and Privacy: (In)visibilities and Infrastructures Springer International Publishing 2017

Objective 5: Engage with the public to demonstrate the benefits of using health data in research and encourage the support of secure and trusted access to patient information.

Throughout 2016 and into 2017, The Farr Institute has continued to position public involvement and engagement at the strategic forefront of its research and communications activity. This section highlights key events and achievements from across the Institute’s Network and regional Centres from April 2016 to March 2017.

Public Panel Activity

The Farr Institute’s four regional public panels act as the public voice in the Institute’s work. Members scrutinise and advise on governance systems, public engagement plans and research practices. This ensures that The Institute’s research is conducted appropriately, securely and transparently and for the benefit of patients and the public.

All four public panels have continued to meet regularly, maturing further in their roles. With support from public engagement staff, members are using their expertise and experience to act as advisors, representatives and collaborators to shape and improve the research of The Farr Institute. This has included the following roles and responsibilities:

- **Critical review.** Offering researchers regular opportunities to meet with panels to act as ‘critical friends’, offering constructive scrutiny on early stage proposals and ongoing research projects. During 2016/17, at least 14 researchers presented to panels about diverse topics including determining reasons for anxiety amongst people living with MS, the use of smart phone data in health research, using routine health records to determine the severity of diabetes, research access to blood spot samples collected at birth (Guthrie cards) and links between mental and physical health.

- **Strategic planning and development.** Collaborating with public engagement staff to co-produce and evaluate public engagement strategies and work plans for individual centres (e.g. CIPHER and Scotland) and collaborative activities (e.g. HeRC and Connected Health Cities).

- **Advising on communication and dissemination.** Members are actively working with our public engagement and communications teams to explore ways of communicating the benefits of health data science to various audiences, including the general public. For example, members of the CIPHER public panel reviewed a presentation produced for a project investigating the influence of antipsychotics on children, which the researcher reported has now improved dissemination of his project findings and generated new work streams.

- **Leadership and collaboration.** Public panel members have been involved in research grant applications, including acting as formally named collaborators. For example, Miss Suzy Bourke from the HeRC panel is named in the ‘Wearable Clinic’ (EPRSC, £1.6 million, PI: N Peek).

- **Conference planning and participation.** Members of all four public panels led a popular panel session at the IPDLN 2016 conference, filling the Swansea University Bay Campus Great Hall. The panel members firstly shared views and experiences of the use of linked data research for the benefit of societal health and then answered questions from delegates covering the topics of; concerns about data privacy, perceived benefits of research, building public trust, and communication. The impact of the session was felt beyond the conference delegates, generating a wealth of positive conversations outside of the conference including on social media. Furthermore, the organising committee for the 2017 Informatics for Health conference includes a public panel member, who has been involved in influencing decisions regarding the programme, communications and delivery.

Citizen Driven Health

The Farr Institute’s Centres have led a number of pioneering research projects in which the public have been at the heart of setting the research agenda, providing the data and analysing the results.

#BritainBreathing

Over 1,500 participants have taken part in #BritainBreathing, a ‘citizen science’ project to investigate the possible environmental causes of rising allergies. The project uses a smartphone app, co-designed by allergy sufferers, to collect data whilst empowering citizens to understand and manage their own symptoms. Public engagement events have been held as part of Pint of Science, the European Science Open Forum, Cheltenham Science Festival and Bluedot, a major new music and science festival, where over 1000 peak flow tests and 250 ‘ink breath’ drawings were completed during the course of the weekend. In addition to The Farr HeRC
Public and Patient Involvement and Engagement

eHRC award the research is supported by BBSRC, British Society of Immunology and the Wellcome Trust.

www.britainbreathing.org

ACTIVE Project
This is a collaboration study between Farr CIPHER and the British Heart Foundation launched in January 2017. The study involves Year 9 pupils (age 13-14 years) at eight randomly selected schools in the Swansea area. Each of the engaged pupils receives £20 per month in vouchers to spend on physical activities of their choice. The study aims to see whether giving teenagers’ vouchers to spend on activities of their choice can reduce the time spent being sedentary, improve fitness, lower risk of heart disease and improve general health.

The pupils are at the heart of the study, they developed the project concept themselves using the results of a previous feasibility study, they have set the study’s research aims and helped the research team direct the way in which ACTIVE runs. The pupils are also currently establishing a social media campaign to promote the ACTIVE project and its results, to make evident to others how physical activity can improve an individual’s long term health and wellbeing. The goals of the study include; encouraging pupils to take ownership of their own health and wellbeing, and act on the evidence base which they have produced, and to work in collaboration with the local council to adapt their service provision to meet the physical activity needs of young people in the local area.

www.publichealthnetwork.cymru/en/good-practice-directory/the-active-project

Cloudy with a Chance of Pain
Launched in January 2016, this project is the world’s first smartphone-based study to investigate the association between arthritic pain and the weather. The study was carried out during 2016 using a smartphone app which people used to record how they are feeling, whilst local weather data was automatically collected using the phone’s GPS. The study was open to anyone in the UK with arthritis or chronic pain and aged over 17. As well as recording their symptoms participants were also able to help researchers analyse and interpret the data collected. The research is supported by Arthritis Research UK, uMotif (London) and the Office for Creative Research (New York). The project recruited over 13,000 participants and now that recruitment is closed researchers are in the process of carrying out a formal analysis of the data.

www.cloudywithachanceofpain.com/the-project

HAPPEN
Health & Attainment of Pupils in a Primary Education Network (HAPPEN) is a network of health, education and research professionals aimed at improving the health, wellbeing and education outcomes of children aged 9-11 in Swansea. Farr CIPHER researchers were able to link and study child health behaviour data, health records and education data. Researchers are using this data to provide feedback reports to the schools, and organisations (dieticians, sport development, local charities, and public health professionals). The reports were written in line with the curriculum framework to ensure the children’s learning is enhanced, and to enable schools to see how they compare with other schools in their county to identify areas of need. On-going support and advice is provided to the schools via the HAPPEN website which gives teaching staff access to resources which they can use to improve health and wellbeing interventions. The use of consultation, engagement and collaboration has enabled the network’s success to date and the number of schools joining the network is continuing to grow. The partnership between schools, health professionals and Farr Institute researchers provides a more unified and evidence based approach to help tackle child health, well-being and education.

www.happenswansea.co.uk

Public Preferences Regarding Data Linkage for Research: A Discrete Choice Experiment
Previous research has identified a range of factors which influence whether the public support data linkage for research. In order to examine the relative importance of factors influencing public preferences a Discrete Choice Experiment (DCE) was conducted via an online questionnaire among members of Ipsos MORI’s online panel in Scotland. The survey was launched on 12 August 2016 and was live for a period of 14 days. In total 1,004 respondents completed the full survey. The results showed that overall the two most influential factors shaping respondents’ preferences are the type of data being linked and how profits are managed and shared. The type of data being linked is roughly twice as important as who the researchers are. There were slight differences across age groups and between genders and slight differences when comparing respondents with and without long term health conditions. The most notable differences were observed in looking at differences in employment and working sector. Respondents who were not working full time were more concerned with oversight arrangements and the type of information being linked, compared to those who were working full-time. Those working full-time were less concerned with oversight arrangements and more concerned with the purpose of data-linkage, who the researchers were and how profits were managed and shared. The outcomes of this study may be useful for indicating which factors to focus on in future public engagement and has important implications for the design and delivery of research and public engagement activities. The continuously evolving nature of the field means it will be necessary to revisit the key conditions for public support on an ongoing basis and to examine the contexts and circumstances in which these might change. (Manuscript in preparation).

Resources for the Publics

100 Ways of Using Data to Make Lives Better
A major communications project for the Institute, this national series of 100 case studies showcases not only the work of The Farr Institute but of the wider UK research community. Through this project, the Institute aims to promote the safe and trusted use of data in research and raise awareness of its benefits to patients and the public, hoping to counter some of the recent bad press around the topic. Highlights of the project include:

- Case studies are published on The Farr Institute website and can be downloaded as PDFs.
- The series is shared on social media as part of the #datasaveslives campaign.
- A communications plan will be produced for 2017 to highlight relevant 100 Ways case studies throughout the year to coordinate with major events and national awareness days.

www.farrinstitute.org
A searchable database will be produced to allow the most appropriate studies to be identified and used as support material for disease or research-specific events and activities.

Some of these case studies have contributed to resources made available by The Understanding Patient Data initiative, led by Wellcome Trust.

Organisations which have shown an interest in using the case studies to promote the use of data to improve public health and care include Scottish Government, Department of Health, Children’s Policy Research Unit (UCL) and MQ Transforming Mental Health.

The case studies featured at the All Party Parliamentary Group Medical Research meeting on 12th July 2016 at the House of Commons.

Other resources produced include:

- 10 cancer patients video stories on why health data is important to health research and service improvement
- #BritainBreathing and Cloudy with a Chance of Pain videos to promote citizen science
- Body Beats video, highlighting the process of turning health data into music
- #datasaveslives Community Reporters’ videos
- ‘Who We Are and What We Do’ promotional leaflet
- ‘Collect Data Once, Use it Often’ posters and banners to demonstrate the benefits of making routinely collected data available for research
- Citizens’ jury feedback video including contributions from Dame Fiona Caldicott

Understanding Patient Data Initiative

The Farr Institute is supporting Understanding Patient Data; an independent initiative led by Wellcome Trust which was established in autumn 2016. The initiative is funded by Wellcome Trust, MRC, Department of Health, ESRC and Public Health England. It will support better conversations about the uses of health and care information, building on the work of the Caldicott review, helping to develop a framework for clear and transparent discussions with the public, patients and healthcare professionals about how data can be used to improve health. It will develop innovative approaches and tools to encourage more effective dialogue and communication. These include; phrasing and terminology guidelines; a database of case studies; curated studies about public attitudes towards the use of data and sharing best practice.

Understanding Patient Data’s official launch and first meeting took place on 6 April 2017.

For more information visit www.understandingpatientdata.org.uk

Website & Social Media
#datasaveslives

- The Farr Institute Network and the four Centres have a combined Twitter following of more than 7,000.
- The Institute continues to highlight the positive impact of health informatics research on public health through this popular social media campaign. Highlights include:
  - Support from AstraZeneca, the Northern Health Science Alliance, MQ, The Stroke Association, World Health Organisation, Brain Tumour Charity, Association Medical Research Charities, the northern Academic Health Science Network, Dame Nancy Rothwell, Dame Fiona Caldicott, George Freeman MP, Jo Churchill MP, Dame Sally Davies, Prof Ian Greer and Dr Bruce Keogh.
  - Adoption of #datasaveslives by the Department of Health’s £20m Connected Health Cities programme www.connectedhealthcities.org.
  - 10 cancer specific #datasaveslives videos were produced in partnership with the Use My Data campaign, led by the National Cancer Intelligence Network (NCIN).

Website

In an effort to unify The Farr Institute brand and to truly represent the nationwide and cross-Centre approach to the Institute’s research and activities, The Farr Institute launched a new website (www.farrinstitute.org) on 13 April 2016.

Creating a public-friendly website accessible to the lay reader was crucial to the brief for this project. While communicating the public benefits of using data in research and articulating how this is done at The Farr Institute, it was important to provide a platform to demonstrate the Institute’s dedication to public opinion and engagement through initiatives such as the #datasaveslives campaign and the 100 Ways of Using Data to Make Lives Better case study series.

From April 2016 to April 2017 www.farrinstitute.org has achieved:

- 24,835 unique users from 134 different countries
- 38,621 user sessions
- 115,792 page views

Dame Sally Davies, Sir Bruce Keogh & Amanda Lamb support #datasaveslives
Events and Activities

The past year has seen a busy events calendar with activities hosted by regional Centres and external events that the Institute has participated in. This has included nine courses and activities for researchers about public engagement, and 30 events and activities which were open to the public. A full list can be found in appendix 4.

Highlights include

- Participation at major national science festivals including, Cheltenham, Manchester, Bluedot and Edinburgh. The Institute was also represented at the British Science Festival in Swansea where ‘Dave the Data item’ and his team used a jigsaw analogy to explain how de-identified data can be linked for research.

- The Cabaret of Dangerous Ideas at the Edinburgh Fringe Festival. Dr Mhairi Aitken asked the audience to consider the statement “Hands Off My Clubcard... Just take my DNA!”.

- Public debates and ‘Café Science’-type events including the #Britain Breathing Pint of Science in Manchester and Sarah Cunningham-Burley’s lecture ‘Who Owns My Genome?’ in Edinburgh.

- Researcher training including social media workshop, advice sessions on involving the public in research and hackathons.

- Victoria Allan, a Farr Institute PhD student at UCL was shortlisted for the MRC 2016 Max Perutz Science Writing Award. Her article entitled ‘Preventing a heart that goes ba-boom, ba-, ba-, ba-, boom, ba-boom’ described her research on risk factors for developing the heart rhythm disorder atrial fibrillation. https://www.mrc.ac.uk/documents/pdf/max-perutz-essays-2016/

- Our Lives in Data: Science Museum Exhibition Launch, July 2016. The Farr Institute was invited to the launch of this exhibition which featured a 3D data visualisation tool developed by Farr Institute researchers at Bristol, intended to help researchers detect patterns in large, complex linked datasets.

Mathematics Masterclasses for Girls

Through the spring term of 2017, Swansea University hosted the Mathematics Masterclasses for Girls, a series of Saturday sessions aimed at female Year 10 pupils who have a keen interest in mathematics. Through a variety of topics, the pupils were given a taste of the nature of mathematical thinking by research staff based at the University with an aim to encourage even more girls to pursue scientific studies and careers.

As part of an on-going programme of public engagement for Farr CIPHER, a team of female researchers and analysts from CIPHER and the Office of National Statistics (ONS) conducted a master class session: ‘The Spread of Disease’. The session focused on how mathematics is used to measure the frequency and spread of diseases by using probabilities associated with the risk of infection.

The Spread of Disease session attracted more than 40 pupils, all of whom engaged in the extremely interactive feel of the session. Pupils were offered the opportunity to explore how hygiene can affect the spread of diseases, mathematics and glitter were used to do this! A simplistic dice game was used to calculate and demonstrate the statistics relating to the spread of the common cold.

By the end of the session, all the pupils were aware of how important mathematics is for exploring patient and healthcare data – to identify trends in disease and treatments. It highlighted that if you love mathematics and have an inquisitive mind, then there is a job for you in health.
Mainstream Press and Media Coverage

As well as features, news and articles in specialist publications, The Farr Institute has gained mainstream public exposure through press and media over the past year. Some examples are highlighted below.

- Gradual increase in antidepressant use among children and adolescents (Linda Wijlaars), March 2016 (BBC, Independent, Huffingon Post, Daily Mail)
- Launch of the #BritainBreathing project (Dr Lamiece Hassan), March 2016 (BBC Radio 4 Inside Science, Daily Mail, The Independent, BBC Click)
- Antidepressant prescribing to children and young people (Ann John), March 2016 (BBC, Good Morning Wales, Wales Online, South Wales Evening Post)
- Failure of the Quality and Outcomes Framework in reducing premature mortality (Dr Evan Kontopantelis), June 2016, (Daily Mail, The Telegraph)
- The impact of the media’s reporting of research into the side-effects of statins in 2013 (Tjeerd van Staa), June 2016 (Telegraph, The Guardian)
- Early results from the Cloudy with a Chance of Pain project (Dr Will Dixon), July 2016 (Daily Mail, i paper, FT, BBC)
- Over 1,000 asthma patients ‘dying needlessly’ every year (Professor Aziz Sheikh), August 2016 (Sky News)
- Adult lifetime weight gain and obesity related cancer (Dr Hannah Lennon), December 2016 (Daily Express, i paper, Daily Mail, The Guardian, The Telegraph, Men’s Health)
- ACTIVE: Combatting teenage inactivity (Sinead Brophy), January 2017 (ITV Wales, BBC Radio Cymru, BBC Radio Wales)
- Schools shouldn’t be left alone to deal with child health and well-being any longer (Emily Marchant) January 2017 (World News Network, Yahoo News)
- ‘Can Tech Save the NHS?’ (Iain Buchan), March 2017 (PC Pro)
- Campaign to reduce the risk of suicide in young men, (Ann John), March 2017 (ITV Wales)
Partnerships

Objective 6: Bring together government, public sector, academia and industry to foster relationships and establish best practices for innovation, discovery and impact in health data science.

The Farr Institute continues to establish collaborations across the globe with the leaders of health data science in other countries. Figure 3 shows current international collaborations established by The Farr Institute.

Partnering with Industry

This year’s focus has been the consolidation and development of existing and strategic mutually beneficially relationships. In addition we continue to extend our reach, with around 100 new leads created through partnership with events including BioData World Congress, BioDundee and DataFest 2017. A coordinated plan of event attendance for lead generation across 2017 is currently in implementation, including a central presence at this year’s Farr Institute international conference: Informatics for Health 2017, Manchester, April 24-26.

The Farr Institute has built on and extended its strategic relationship with Intel’s Heath and Life Sciences team further over the course of 2016, holding a collaborative workshop in Scotland to explore the challenges arising from creating a national medical image data set. The Institute has also commenced a collaborative project, co-funded with the Alan Turing Institute and Intel: Federated Access to Secure Healthcare Information, as outlined in the Enabling datasets segment of this report (page 9).

In addition, Intel are working with the Farr HeRC team on a collaborative research project to advance technology to improve treatment of pressure ulcers. The partnership
also includes GPC Limited and Pennine Care’s NHS FT, working together using the Connected Health Cities care pathway framework to identify the optimal technical solution for introducing electronic wound care management in community settings. Pilot testing is scheduled to commence in April 2017 in collaboration with a dedicated Wound Care service in one of the Greater Manchester boroughs.

Farr researchers (T van Staa, I Buchan et al) have been instrumental in driving a strategic relationship between the Greater Manchester city region and Janssen plc, to develop a suite of industry-funded research projects into treatment resistant depression. These are aligned to the Greater Manchester devolution agenda and include digital health, informatics and clinical trials - details subject to signed confidentiality agreements at present. Workshops and focused meetings have now led to a stage of more defined programme planning for 2017 involving a multi-disciplinary team, including clinical psychologists (S Bucci), psychiatrists (J Deakin, S Lewis), informaticians and trialists (I Buchan, T van Staa) and software engineers (M Machin, P Whelan).

Many large pharma companies including Bristol Myers-Squibb, Pfizer and Sanofi have made approaches regarding the availability of data, linkage and analysis across The Farr Institute and commercial in confidence discussions are currently underway. Monitoring of these opportunities will be managed through the central customer relationship management system to enable scaling where appropriate.

**Multi-partner Large Scale Collaborations**

In Wales, this work is most pronounced in the ARCH (a Regional Collaboration for Health) consortium with enhanced multi-sectoral data linkage on a population of 1 million in South West Wales being used to drive the case for change in modernising health services, including the development and evaluation of a Wellbeing Programme. The Farr Institute and the partner National Centre for Population Health and Wellbeing Research have been commissioned to contribute to the development of the business plans. Related to this development is the Swansea Bay City Region deal backed by the Chancellor in the Autumn Statement. The ambition is to create an ‘internet coast’ turning the whole of the region into a digital superhub to transform the regional economy, the future of energy and the future delivery of health and social care. There are also plans for a new transatlantic fibreoptic cable from New York to Swansea and a 5G test bed that would hugely increase capacity for high volume health analytics.

Building on the expertise in cardiovascular science and collaborative activities and proposals undertaken last year, Farr London was awarded £20 million from the Innovative Medicines Initiative. The research programme, BigData@Heart, will integrate healthcare data, wearables, genetic profiles, information about patients’ lifestyles and health and their own reporting of symptoms, to better understand the causes of these conditions and the different subtypes. The BigData@Heart consortium includes the European Society of Cardiology, European patient organisations, universities from Utrecht, Berlin, London, Cambridge, Valencia, Stockholm, Hamburg, Birmingham and Uppsala and various pharmaceutical and technology partners, allowing for rapid technology transfer and implementation of research outcomes.

The Spin-in Lab is a new innovation environment to build tri-partite partnerships between industry, NHS, and the University of Manchester to accelerate the development of digital health technologies. This model enables companies to develop, test and validate digital health products and services using real-world health data and access to domain specific data analytics capabilities and expert advice on the design and development of their ideas. The most recent Spin-in Lab project is a Farr HerC collaboration between Intelligent Medical Objects (IMO), Salford Royal NHS Foundation Trust and the University of Manchester to support a research and innovation project to advance the automation of coding diagnoses from digital clinical text. The focus of this work is on retrospective coding of a rich collection of clinical data in order to:

- Obtain deeper digital phenotyping for clinical epidemiology
- Enable health systems to learn from structured data to adapt of clinical workflows
- Increase the coding of health issues to demonstrate the advantages of a well-coded patient record

The project’s first sets of results are expected in April 2017. As a direct result of this Farr Institute - Industry collaboration model, IMO will be a strategic partner and will hold a sponsored focus group at The Farr Institute’s next annual conference: Informatics for Health 2017.

Farr Institute investment has established the Prudent Healthcare Intelligence Hub (PHIH), an academic/NHS partnership to improve NHS services. Under this model, NHS organisations across Wales identify the key research questions they need answered in order to improve their services. A team of co-located Farr CIPHER and NHS analysts then utilise the SAIL Databank to provide these answers. This NHS-led approach ensures the research generated has maximum impact on services and patients. Partnering with academia also increases the NHS’s own capacity to link and use electronic data (supported by bursaries provided by Farr CIPHER for NHS staff to undertake Swansea University’s Health Data Science MSc course). This model ensures that data linkage, The Farr Institute and SAIL Databank become an increasingly important part of NHS organisations’ day-to-day business intelligence, commissioning and service improvement activities. In the first year of operation, nine projects have been developed with Abertawe Bro Morgannwg, Aneurin Bevan, Hywel Dda and Cwm Taf University Health Boards, the Welsh Ambulance Service and Public Health Wales NHS Trusts. These projects cover a wide range of activities and diseases areas including physical activity, injuries, cancer, and cardiovascular disorders.
Outcomes from the IPDLN Conference include

New funding. Presentation (Bibliometric Analysis to Scan and Scrape new datasets: It’s all about that BASS) at IPDLN by K Tingay and A Anastasiou, resulted in 2 new grants.

- Scientometrics: European Medical Information Framework (EMIF): funding of £119,227 to automate queries to identify data sources from publications in order to meet EMIF and Innovative Medicines Initiative (IM2) project milestones.
- Family identifier: Cardiff University Seedcorn Funding of £4,370 to develop a family identifier using SAIL data, which can be used to speed up analyses on mother-child effects on health and wellbeing.

New papers. The outcomes of the workshop, Entities and identities, led by Kerina Jones are informing the consensus definition of the field of population data science, to be published shortly in a position paper.

Recruitment into MSc Courses. Following and as a direct result of IPDLN, six delegates (from UK, Brazil and Canada) were recruited and went on to attend the MSc Health Informatics courses at Farr CIPHER, Swansea University and five delegates from Canada and Australia have signed up for a webinar.

New Collaborations.

- Presentation by Farr CIPHER researchers entitled ‘Codifying unstructured data: A Natural Language Processing approach to extract rich data from clinical letters’ resulted in a new collaboration and a Farr-wide working group has been set up with Manchester and London with an aim to pull together a workshop on Natural Language Processing.
- A secondment of a Welsh Government official to Swansea University (and a place on its Health Data Science MSc course), to use SAIL Databank data to carry out policy-related research and increase the Welsh Government’s capacity in data linkage research.
- As a direct result of the IPDLN ‘Exploratory data analysis using R as a GIS’ workshop and ‘Geocoding routinely collected administrative data to measure access to alcohol outlets in Wales’ presentation, Richard Fry is working with Sarah Mah from McGill University, Canada looking at new research around neighbourhood walkability and health outcomes.
- A team from Farr CIPHER led by Ashley Akbari is working with colleagues from Farr Scotland on a validation project titled TSOAC (The risk of death and all Serious vascular events for patients who start a non-vitamin K Oral Anticoagulant after intracranial haemorrhage), see page 9.

Formalising a Partnership with the Institute of Clinical Evaluative Sciences, Ontario (ICES)

In July 2016, The Farr Institute signed a Memorandum of Understanding (MOU) with the Institute for Clinical Evaluative Sciences (ICES), solidifying a commitment to work together in areas of common strategic operations and interest.

ICES is a not-for-profit community of research, data and clinical experts based in Ontario, Canada which leads cutting-edge studies that evaluate healthcare delivery and outcomes. Both ICES and The Farr Institute share the objective of being world-leading research institutes delivering research excellence that results in trusted evidence which in turn makes policy better, healthcare stronger and people healthier.

Noting that a number of informal collaborations between the two institutes already exist and after exploring common aims and interests, ICES and The Farr Institute decided to establish the MOU to set out areas where both institutes might work more closely together. By pooling expertise and resources, the partnership will contribute to the international advancement of outcomes in areas of health data science including:

- Capacity building
- Methodology, data science and research
- Public involvement and engagement
- Governance, policies and procedures
- Research exemplars

ICES and The Farr Institute met for the first time since the partnership was formalised at the IPDLN Conference in Swansea, Wales in August 2016. Leads from each institute were identified to form a working group which will develop the collaboration by identifying shared best practices, potential collaborative research projects and opportunities to become more closely involved in each other’s programmes. The working group has established terms of reference and is holding quarterly meetings to develop shared activities and dissemination of best practice.

Partnering to Expand the Reach of The Farr Institute International Conference

In 2016 and 2017, The Farr Institute has partnered with other international organisations to host the annual conference. This has allowed the Institute to increase the number of delegates attending these events and broaden the global reach and impact of The Farr Institute.

As its Director, Prof David Ford (Deputy Director Farr CIPHER) hosted the International Population Data Linkage Network biannual conference in Swansea in August 2016. The Farr Institute co-hosted this event along with the Administrative Data Research Network. The three day conference saw a record number of 500 attendees as international data linkage professionals descended on Swansea from 22 countries, spanning six continents. This created many opportunities to meet new peers, find out about the fantastic work being conducted using data linkage from around the world and open up the potential for new collaborations.
In April 2017 The Farr Institute is joining with the European Federation for Medical Informatics (EFMI) to bring together The Farr Institute annual conference and Medical Informatics Europe under the banner of Informatics for Health 2017. EFMI is a non-profit organisation concerned with the theory and practice of information science and technology within health and health science in a European context and represents 32 countries. At time of writing, registrations exceed 700; 290 gala dinner seats have been booked; and sponsorships include 1 gold, 1 silver and 7 bronze. The extensive and varied scientific programme is complemented by a programme of social, ‘art-meets-science’ and public engagement events including a play (The Nest), ‘Pint of Science’ (Alan Tucker) and the innovative ‘Science Slam’. The Science Slam aims to help researchers deliver their science in an accessible format to non-specialists. Participants have 8 minutes to present their research and ideas in an entertaining or unusual way. The audience forms the jury and decides the winner.

The annual Farr Institute Doctoral Symposium will be run as one of the pre-conference events and has been opened to all PhD students attending the conference. Attendees will have the opportunity to present their doctoral research and, under the guidance of senior academics, identify training needs to further their careers in data science.
Appendices
Appendix 1: Publications

A list of publications from The Farr Institute Centres April 2016-March 2017. Publications result from research either directly funded by the original grants (eHIRC and capital spend) or through use of resources funded by the grants with additional funding leveraged from the creation of the Farr Institute Centres.


e31007a4-c0c4-428a-8a70-d99bee70284e


Appendices


Fry, R; Morgan, J; Rodgers, SE; Orford, S; Fone D. Using routinely collected administrative data in public health research: geocoding alcohol outlet data. Appl Spat Anal Policy. 2016;


Informatics [Internet]. 2016;23(3):847. Available from: https://www.jihi.org/content/early/2016/01/09/bjp.bpi.114.158932.abstract


Robertson ARR, David; Giunchiglia, Fausto; Pavis, Stephen; Turra, Ettore; Bella, Gabo; Elliot, Elizabeth; Morris, Andrew; Atkinson, Malcolm; McAllister, Gordon; Manatkit, Areti; Papapanagioultos, Petros; Parsons M. Healthcare data safe havens: towards a logical architecture and experiment automation. J Eng [Internet]. 2016; Available from: http://bjgp.org/content/67/555/e86.abstract


Royston P, Parmar MKB. Augmenting the logrank test in the design of clinical trials in which non-proportional hazards of the treatment effect may be anticipated. BMC Med Res Methodol. 2016;16(16).


Royston P, Parmar MKB. Augmenting the logrank test in the design of clinical trials in which non-proportional hazards of the treatment effect may be anticipated. BMC Med Res Methodol. 2016;16(16).


Appendices

Turner M, Barber M, Dodds H, Dennis M, Langhorne P, Macleod M-J. Stroke patients admitted within normal working hours are more likely to achieve process standards and to have better outcomes. J Neurol Neurosurg Psychiatry [Internet]. 2016;87(2):138 LP-143. Available from: http://jnnp.bmj.com/content/87/2/138.abstract


Yebra G, Hodcroft EB, Ragonnet-Cronin ML, Pillay D, Brown AJL. Using nearly full-genome HIV sequence data improves phylogeny reconstruction in a simulated epidemic. Sci Rep [Internet]. The Author(s); 2016;6:39489. Available from: http://dx.doi.org/10.1038/srep39489


Appendix 2: Markers of esteem

Awards

**Ann Blandford**
Suffrage Science, MRC Clinical Sciences Centre. The award celebrates women in science for their scientific achievement and for their ability to inspire others.

**Iain Buchan**
Named as one of 14 'Mbassadors' (high profile industry ambassadors for Manchester city) by Marketing Manchester, in recognition of Iain's contribution to the life science sector.

**Mark Caulfield**
Bjorn Folkow Award of the European Society of Hypertension.

Ranked amongst the world’s 200 Most Highly Cited researchers in the field of ‘Molecular Biology & Genetics’ by Thomson Reuters based upon ranking in the top 1% for citations in the field and year of publication.

**Helen Colhoun**
AXA personal Chair in Medical informatics and Life Course Epidemiology, University of Edinburgh. In recognition of her exceptional expertise in the use of electronic health data to model the risk of diabetes, and her vision to develop digital health tools for individual preventive strategies and for precision medicine.

**Paul Foster**
Fellowship by Election for services to surgical science 2016, Royal College of Surgeons of England.

**John Frank**
Visiting Professorship, University of Western Australia/ Western Australian Health Transfer Network. In recognition of his long experience, in Canada and the UK, in building research collaborations with research users in the public health and health systems policy and practice communities - including his knowledge of various models of governance for cross-sectoral record linkage for research.

**Benjamin Green and Guy Jackson**
Public Health England organised hackathon, the event was held in Manchester in May 2016 and was designed to help generate new ideas for a series of digital products designed to tackle childhood obesity. The final ideas are top secret and are currently being developed by PHE.

**Philip Hannaford**
Fellowship Royal Society of Edinburgh, for distinction in epidemiological research particularly that related to safety of contraceptives.

**Lamiece Hassan and #BritainBreathing team**
Better World Award 2017, Faculty of Science Engineering, The University of Manchester, The #BritainBreathing project in the category of outstanding contribution to public engagement.

**Harry Hemingway**
NIHR Senior Investigator.

**Aroon Hingorani**
NIHR Senior Investigator.

**Marjory Johnston**
Young Researchers prize for oral presentation at 9th European Public Health Conference (Vienna), November 2016. First (joint) prize, as an early career researcher for presenting her work "The influence of educational attainment on the association between social class at birth and multimorbidity in middle age in the Aberdeen Children of the 1950s birth cohort’ at World Congress on Public Health, Melbourne, April 2017.

**Clemence Leyrat**

**Jill Pell**
CBE for services to public health research. Jan 2017.

**Richard Williams**
British Computer Society’s John Perry prize, work with the Patient Safety Centre on the SMASH project was "Highly Commended"

**Rebecca (Becca) Wilson**
Fellow of the Software Sustainability Institute (SSI) for demonstration of a sound understanding of the rigorous approach that must be adopted if software development is to be professionalised and sustainable.

Memberships

**John Ainsworth**
Manchester Informatics Executive Board; Advisory, Journal of Mobile and Ubiquitous Health - Editorial Board; The Manchester Connected Health Ecosystem; School of Health Sciences Research Committee.

**Ashley Akbari**

**Michael Barnes**
MRC Methodology Research Panel; MRC Stratified Medicine Expert Panel; CORBEL Open Call scientific review panel; Chair, Innovative Medicines Initiative eTRIKS knowledge management, Science Advisory Board.

**Paul Burton**
Expert Advisory Group for Data Access to MRC, Wellcome Trust, ESRC and CRUK.

**Marion Bennie**
Member of Control of Antimicrobial Resistance in Scotland (Policy Group); Country representative for Scotland of the European Drug Utilisation Research Group - Executive Committee; Member of Scottish Antimicrobial Strategy Group.
Appendices

Ann Blandford
Chair, UK Computing Research Committee.

Iain Buchan
Chair, Connected Health Cities Advisory Board. Member of: MRC Methodology Advisory Group; Health Innovation Manchester Board and the GM IM&T Leaders Group; NHS Digital Health Academy steering group; GM Connect Executive Steering Board, Manchester; MRP Advisory Group, London; CHESS Steering Committee Meeting, Salford; NIB Working Group; GM CHC Steering Group.

Gwyneth Davies
Asthma UK Cymru; Respiratory Health Implementation Group (Wales).

John Deanfield
Senior Advisor to PHE, Cardiovascular Disease Prevention, Governance Dementia Board, Brain Age Project, Health Check Expert Scientific and Clinical Panel (ESCAP); CV Academic Lead, National Cardiovascular Intelligence Network (NCVIN); Academic Chair, Heart UK PASS Governance Committee; Advisor to NHS England, Familial Hypercholesterolaemia Task and Finish Group, Healthchecks Heart Age Working Group and Healthchecks Working National Evaluation Group.

David Ford
Panel expert, Research Council of Norway; Committee member for; HealthWise Wales Scientific Steering Group; EDARA; Health Informatics task force; ELASTIC.

Paul Foster
UK Dementias Platform (UKDP) Dementias Readiness Cohort Imaging Working Group; Faculty, AAO Pyott Glaucoma Education Center; NICE Glaucoma Guidelines Committee; Royal College of Ophthalmologists Academic Sub-Committee of Scientific Committee; Member, Executive Committee – UCL Institute of Ophthalmology.

John Frank
Chair of Advisory Committee to Quasi-Experimental Evaluation of the Family Nurse Partnership in Scotland (Scottish Government contract won by a Welsh consortium); Chair of International Scientific Advisory Committee, OPTIMISE Project (CIHR Grant awarded to Prof Laura Rosella et al., University of Toronto, to improve the use of routinely collected PPH data for health system and policy action); Member of Grant Review Panel for Public Health, Agence Nationale de Recherche (France), Member of Evidence and Data Sub-Group, Scottish Government Strategic Review of Maternity and Neonatal Services; Member of Scientific Advisory Group, Better Start Bradford Innovation Hub.

Ruth Gilbert
UK Biobank Board.

Philip Hannaford
Member of WHO Expert Group on Review of Evidence of HIV acquisition and hormonal contraception; Member of NIHR Research Professorships selection panel.

Harry Hemingway
Advisory Board member, German Hospital Informatics Research Consortia; Journal editorial committee for, BMC Cardiovascular Disorders; Heart; Circulation: Quality and Outcomes Research; European Heart Journal: Quality of Care and Clinical Outcomes; Lead Academic for Biobank EHR phenotyping for cardiac outcomes working group; Founder Board Member, National Institute for Cardiovascular Outcomes Research (NICOR); Founder Board Member, Myocardial Infarction National Audit Project (MINAP) Academic Group; NICE Observational Data Expert Group; Lead the Farr Electronic Health Record Clinical Interpretation Partnership, Genomics England 100,000 genomes project; Office for Strategic Coordination of Health Research (OSCHR) – Informatics Sub-Group.

Andrew Morris
Wellcome Trust/MRC/ESRC/Cancer Research UK Expert Advisory Group on Data Access (EAGDA); Chairman Scottish Government Data Linkage Framework Board; Member of the Life Sciences Advisory Board for Scotland; Governor of the Health Foundation; Chair, Wellcome Trust Investigator Awards Expert Review Group, Genetics, Genomics and Population Research; Member Board for Academic Medicine; Member MRC Stratified Medicine Steering Group and grant panel; Chair UCL Partners Informatics Programme Board; Chief Scientist (Health) for Scotland; Member Scottish Government 2020 Vision Board; Member of the Office Strategic Coordination of Health Research (OSCHR) Board.

Jill Pell
Member of BHF Chairs & Programme Grants Committee; Member of Cancer Research UK Epidemiology Expert Review Panel; Member of CRUK, Wellcome Trust & BHF, Population Health Working Group.

Steffen Petersen
Chair-elect, Cardiovascular MRI; European Association of Cardiovascular Imaging (EACVI); European Society of Cardiology (ESC).

Alan Radford
Advisory, Member of the British Veterinary Association Surveillance working group.

Tjeerd van Staa
HQIP Methodology Group; HRA National Research & Ethics Advisors’ Panel (NREAP), NICE Observational Data Expert Group; Steering Committee of study of preterm birth and neuropsychiatric genes (SPRING); Steering Committee of ERA study: Examining the challenges, opportunities and workforce implications of Electronic Records in Ambulances.

External Meetings - Invitations to Speak and Participation on Panels

Mahmood Adil, Sarah Gorst and Marie Kane
Invited speakers to Evidence Europe 2017, London February 2017

John Ainsworth

Invited Speaker DigitalHealth UK: Transforming Healthcare delivery, Manchester, October 2016

Mohammad Al Sallakh
‘Correcting the annual prevalence of treated asthma for incomplete GP registrations’ IPDLC, Swansea, UK.

‘Early predictors of quiescent asthma in late childhood’ Welsh Thoracic Society Meeting, Cardiff, UK.

‘Early predictors of quiescent asthma in late childhood’ ERS Congress, London, UK.
Appendices

‘Concordance between self-reported and GP-reported treated asthma in Wales’ PCRS-UK Conference, Telford, UK.

‘Making sense of patient-reported currently treated asthma using routinely collected data’ BTS Winter meeting, London, UK.

Michael Barnes
‘Translational Bioinformatics’ Penn Transplant Institute, University of Pennsylvania – invited seminar, Philadelphia, USA.

Marion Bennie
‘Utilisation of connected Electronic Health Datasets’ at Centre for Innovation in Regulatory Science: Real World Data to Real World Evidence for assessing efficacy and effectiveness: Opportunities and challenges for new medicines development, regulatory review and health technology assessment, Washington DC.

Damon Berridge
‘Misinformation versus Information: The Statistical Modelling of Health Data’ RSS International Conference, Manchester.


Ann Blandford
‘What can we learn from NRLS? Exploratory studies in interpreting free-text descriptions from incident reports’ 2017 MHRA NHS Improvement event for Medical Devices Safety Officers (MDSOs) and Medication Safety Officers (MOSs), London.

Iain Buchan
Connected Healthcare: Leveraging data and technology to improve health and social care panel, Pharma Integrates 2016, London

HealTex Launch Event, Manchester

Interoperability Summit Event, London

Statistical Horizons: Farr Manchester and BSU Cambridge workshop, Manchester

Scientific Leaders Team meeting and NHS Digital workshop

MMPathIC - Engaging Engineers and Physical Scientists Workshop, Manchester

A workshop to discuss a University Institute of Health Services Research

Greater Manchester Connected Health Cities Industry Workshop, Manchester.


Merck Sharpe & Dohme Corp. Invited Speaker New Jersey, USA, April 2016

‘“Trust me, I am data” - can data sharing become the trusted foundation of future health?’ Panel member with Lamiese Hassan and Tjeerd van Staa, EuroScience Open Forum, Manchester, July 2016.

‘Informatics for Health needs Informatics of Place’ Invited keynote, EFMI HEC 2016, Munich, August 2016.

Paul Burton
‘DataSHIELD: taking the analysis to the data not the data to the analysis’ Workshop on new research developments in techniques for data privacy. Isaac Newton Institute, Cambridge. December 2016.

‘DataSHIELD: taking the analysis to the data not the data to the analysis.’ Annual conference of Doctoral School of Public Health in Paris: Epidemiology and biomedical information sciences – conference on cohort studies. St Malo, October 2016


‘DataSHIELD development’ University of Adelaide, Adelaide.


Invited chair of discussion session, Working Group in Data Privacy and Trust. Research Data Alliance, Denver, US.

‘DataSHIELD: taking the analysis to the data not the data to the analysis’ UCL Farr Institute, London.

‘Flexible disclosure control for microdata using DataSHIELD’ CLOSER Knowledge Exchange Workshop on Disclosure Control, London.

James Carpenter
‘Methodological challenges in observational studies - current perspectives and future directions in functional form’ measurement error and causal inference, STRATOS (Strengthening the Analytical Thinking for Observational Studies) initiative, Banff, Canada.

‘Joint modellling of observational data symposium’ Hasselt, Belgium.

Tom Clemens
“Exposure to ambient air pollution and fetal growth in north east Scotland - a whole population study” American Association of Geographers International Conference 2016.

Adnan Custovic
‘The many faces of asthma in childhood and adolescence: Can we personalise treatment?’ Ann Woolcock Lecture, Woolcock Institute of Medical Research, Sydney, Australia. October 2016.

‘The role of allergy biomarkers in the diagnosis and management of severe asthma’ and ‘Environmental control in the management of asthma: What works, what does not’ Plenary Lectures at 12th International course on paediatric pulmonology, French Riviera. March 2016.
Appendices

Spiros Denaxas
Invited lecture, Transatlantic Health Data Science Workshop, National Institutes of Health, Bethesda, USA.

Invited lecture, University Medical Center, Utrecht, Netherlands.

Invited lecture, Institute of Collective Health, Federal University of Brazil, Salvador, Brazil.

Invited lecture, University of Oxford, Institute of Biomedical Engineering, Oxford.


Invited visit, KwaZulu-Natal Research Institute for Tuberculosis and HIV, KwaZulu-Natal, South Africa.

Invited visit, Africa Centre for Population Health, KwaZulu-Natal, South Africa.

Invited visit, Federal University of Brazil, Salvador, Brazil.

Invited visit, University of Pennsylvania, Institute of Biomedical Informatics, Philadelphia, USA.

Invited visit, Oswaldo Cruz Foundation of Bahia, Salvador, Brazil.

Will Dixon
Cloudy with a Chance of Pain, British Science Festival.

Stephen Evans & Elizabeth Williamson

David Ford
‘How should we use big data?’ Grand Challenge into the cause of Chronic Fatigue Syndrome/ME 2, Bristol.

‘Health Informatic research - Secure Anonymised Information Linkage (SAIL) databank’ Open Data in Healthcare (SCATA), Birmingham.

‘SAIL Databank’ HRB Data Project, Dublin.

‘Safe data linking and research environment’ Meet new DVLA CTO, Swansea.

‘Linking Clinical, Governmental and Patient data: Understanding determinants and outcomes - case studies from Wales & the UK’ ICES, Toronto.

‘SAIL Perspective’ ICES, Toronto.

‘Wales: deep data for research and precision medicine’ Visit by Matthew Speers, MD Life Science Organisation, Dept. for International Trade, Swansea.


John Frank
‘Why the British Birth Cohorts are Unique’ at 70 Birthday Celebrations, 1946 British Birth Cohort, London and Manchester.


‘Best Investments for Health Equity Over the Life Course: Scotland and rUK Compared.’ The W. H. Duncan Memorial Lecture, University of Liverpool, Liverpool.

‘Best Investments for Health Equity in Early Life: Scotland and rUK Compared.’ International Network for Research into Inequalities in Child Health Annual Conference, Barcelona.

‘Priorities for Chronic Disease Prevention: The Case for Upstream Prevention.’ Canadian Institutes of Health Research, Toronto.


‘Best Early Life Investments for Health Equity over the Life Course.’ University College Dublin Childhood and Human Development Research Centre Launch Conference, Dublin.

Harry Hemingway
British Heart Foundation and Alan Turing Institute Exploratory Meeting, London.

‘What are big data and what can we use them for?’ Summer School - Big Data in Health Research, Utrecht, Netherlands.

‘Big data for health: first lecture in new stream on informatics’ European Society of Cardiology Congress 2016, Rome, Italy.

‘Big health record data and big knowledge: emerging approaches applied to AF; Integrating new approaches to atrial fibrillation in an era of digital evolution’ The 6th joint consensus conference of the German Atrial Fibrillation NETwork (AFNET) and the European Heart Rhythm Association (EHRA), Sophia Antipolis, France.

‘How can primary care electronic health records add to phenotyping and discovery efforts in UK Biobank?’ Genome Resources Consortium, Boston, MA, USA.

‘Big electronic health records: some challenges and opportunities’ Big Data Excellence in UK and Germany, Berlin, Germany.


Aroon Hingorani
‘From genes to medicines’ Cardiovascular Seminar Series, Edinburgh.


Clemence Leyrat
‘Propensity score analysis with partially observed confounders: how to use multiple imputation?’ Causal Inference Meeting – oral presentation, London.
Appendices

‘Propensity score analysis with partially observed confounders: how to use multiple imputation?’ International Society for Clinical Biostatistics – oral presentation, Birmingham.

Nathan Lea

Ronan Lyons
‘Harnessing the potential of mental health data science’ MQ Transforming Mental Health Meeting, Swansea, March 2016.


The European Injury Data Base (IDB), EUPHA meeting, Vienna.

‘Socioeconomic inequalities and injury risks: opportunities for using IDB system’ BRIDGE Health EU meeting, Vienna.

‘Influencing wellbeing at the population level’ A Regional Collaboration for Health (ARCH), Swansea.

‘Data science: solving the grand challenge of interdisciplinary working for UK biomedical, clinical and epidemiological sciences’ UK Roadmap for Biomedical Metabolic Phenotyping retreat, Wellcome Genome Campus, Hinxton.

‘Using long bone fractures as an injury incidence indicator’ World Injury Prevention Conference (Safety 2016), Finland.

‘Injury Prevention as Social Change’ World Injury Prevention Conference (Safety 2016), Finland.

International Collaborative Effort on Injury Statistics and Methods, Grand research Challenge – Update, Helsinki.

‘Secure Anonymised Information Linkage (SAIL) system’ Educational presentation to the Confidentiality Advisory Group, Health Research Authority, Swansea.


‘MED-BIO: Links to Farr Institute and CIPHER’ MED-BIO MRC bioinformatics centre, London.

‘Multi-sectoral data linkage for intervention and policy evaluation’ 7th ESRC Research Methods Festival: advancing social science research methods, Bath.

‘What SAIL data tells us about housing and health?’ ABMU Changing for the Better Housing and Health Group, Swansea.


‘Use of data linkage in trauma and emergency care research’ Wales International Seminar on Emergency Medicine, Cardiff.

‘Secure Anonymised Information Linkage (SAIL) and Wales Electronic Cohort for Children (WECC)’ Early Years NCPHWR day, Cardiff.

Colin McCowan

Andrew Morris

‘Healthcare in the Information Age’ UBC 100 Centennial Sessions – Symposium on Health Informatics, University British Columbia. 5th April 2016.

‘MRC Farr and ushering in the future of health informatics’ Generation Scotland Symposium IGGM Western General Hospital, Edinburgh, 6th May 2016


European Commission Conference on Personalised Medicine, Brussels. 1st June 2016.


‘Options and Opportunities for harnessing data science in health research’ National Conference for Clinical Research, Kuala Lumpur. 27th-28th July 2016.

‘Medicine in the Information Age’ Newcastle University Centre for Health and Bioinformatics (CHaBi), Symposium Academy of Medical Sciences Lecture. 12th-13th September 2016.

‘Options and Opportunities for Health Data Science in the UK’ BigData World Congress, Sanger Institute, Hinxton, Cambridge. 26th–27th October 2016.


Data Linkage Scotland Showcase jointly-hosted by ADRC-S, The Farr Institute and Urban Big Data Centre, Technology & Innovation Centre, Strathclyde University, Glasgow. 29th November 2016.

Keynote Address, A Digital Frontiers Meeting, Liverpool Town Hall, Liverpool.

Max Parmar

‘Augmenting the logrank test in the design of clinical trials in which non-proportional hazards of the treatment effect may be anticipated’ INSERM symposium on statistical methods in clinical oncology France, Paris, France.

Niels Peek
‘Analytical Challenges for Smarter Health Systems’ 30th AAAI Conference on Artificial Intelligence, Phoenix, Arizona.
Appendices

‘Three Controversies in Data Science for Medicine and Healthcare’ European Data Science Conference, Luxembourg.


Invited speaker in session ‘Big Data in the NHS-improving care, saving money and increasing research potential’ Royal College Physicians Annual Conference 2016

Kathy Pritchard-Jones
‘Childhood cancer outcomes research: Opportunities for integrating clinical and population-based cancer registry efforts’ Scientific meeting and general Assembly of the European Network of Cancer Registries, Baveno, Italy.


‘Developing new organisational systems to deliver outcomes-based cancer care’ Radboud University Academic Medical Centre conference – New Frontiers in Cancer Research: bridging the gap between biology and daily practice, Nijmegen, Netherlands.

Matt Sperrin
Delivered Keynote session, Survival Analysis for Junior Researchers, UK.

Chair, Panel session with Farr Institute members Evan Kontopantelis Iain Buchan, Damon Berridge, Royal Statistical Society 2016 International Conference, Manchester, September 2016.

Rachel Reeves

‘Describing the epidemiology of Respiratory Syncytial Virus in infants and young children in England’ Farr International Conference, Swansea.

Sarah Rodgers
‘Alcohol Outlet Density And Hospital Admissions For Alcohol-Related Injury: An Electronic Record-Linked Cohort Study’ Safety 2016 World Conference, (International conference), Finland.

Administrative Data Research Network National Conference (June, 2016), London.

Environment and health brokerage event, Improving Health with Environmental Data, Edinburgh.


Liam Smeeth


‘Statins, muscle pain and the media’ Clinical Trials Service Unit, Oxford. 2016.

‘Health data: better care or privacy nightmare’ presentation and panel discussion, British Library Talk Science event, London. 2016.

‘Your health data: really private and really important’ presentation at Science Museum Lates series-Big data, London. 2016.


Adam Timmis
‘Key performance indicators for healthcare systems’ LMCmed Forum, Vienna, Austria.

‘Enhancing electronic communication and clinical data access across primary and secondary care’ Farr London Seminar series, London.

‘Electronic Health Records for CV Event Analysis in T2D’ Oxford Centre for Diabetes, Endocrinology & Metabolism OCDEM, Oxford.

‘Cardiovascular registry research’ London School of Economics seminar series, London.

David Weller

Elizabeth Williamson
‘Propensity scores – over-hyped or under-used?’ International Society for Pharmacoepidemiology mid-year meeting, London.

Shang-Ming Zhou
‘Machine Learning Techniques to Identify and Evaluate Interactive Risk Factors from Complex Epidemiological Data’ International Symposium on Embracing the Internet of Things to Data-Driven Decisions, Manchester, UK.

Appendix 3: Leveraged new grant funding

I. Abubakar
‘Early detection and Integrated Management of Tuberculosis in Europe: E-DETECT TB’ £3.2 million awarded to London from EU CHAFEA.

D. Adlam
‘Cardio-oncology: A high resolution national electronic health record investigation of the interplay between cancer and heart disease 2017-22’ (2017-2022) £1.5 million awarded to London from CRUK/BHF.

T. Avery (PI, Nottingham University)
‘Avoiding patient harm through the application of patient safety indicators in UK general practices (Protect)’ (2017-2020). £2.4 million awarded to University of Manchester (£500,000 of which was awarded to the Farr Institute) from NIHR PGfAR.

M. Barreto
‘Newton International Fellowship: Treating heterogeneity and uncertainty in data integration: study on Brazilian healthcare databases’ £65,000 awarded to London from Royal Society.

M. Barnes
‘Positioning dRugs for iMmune Inflammatory diSEases (PROMISE)’ awarded to London from MRC.

M. Bennie

M. Brophy
‘Pharmacoepidemiological surveillance for children with chronic conditions using data mining’ (Oct 2016-Sep 2017). £15,000 awarded to CIPHER from MRC/Farr Institute (£15,000 of which was awarded to the Farr Institute).

‘RCT of ACTIVE - Randomised Control Trial of Active Children through Individual Vouchers Evaluation’ (Jun 2016-May 2018). £299,904 awarded to CIPHER from British Heart Foundation (£271,087 of which was awarded to the Farr Institute).

S. Brophy & S-M. Zhou
‘A real time pharmacoepidemiological surveillance system for children with asthma (PhD Studentship)’ (2016-2019). £56,382 awarded to CIPHER from Swansea University & Université Joseph Fourier, Grenoble (£56,382 of which was awarded to the Farr Institute).

I. Bruce
‘Manchester BRC’ (Apr 2017-Apr 2022). £25 million awarded to University of Manchester from NIHR (£2.4 million of which was awarded to the Farr Institute).

S. Bucci
‘Actissit II’ (TBC-TBD). £1.3 million awarded to University of Manchester MRC (DPFS).

J. Breuer
‘Why do norovirus pandemics occur and how can we control them?’ (Nov 2016-Oct 2020). £2.6 million awarded to London from Wellcome Trust.

P. Burton
‘World Universities Network Research Development Fund - Workshop on sharing sensitive data’ (Aug 2016). £25,028 awarded to CIPHER from Worldwide Universities Network (WUN) (£25,028 of which was awarded to the Farr Institute).

CLOSER Project (Jan 2016- July 2017). Total funding £350,000 award to University of Bristol £140,000 from ESRC.

S. Campbell
‘PSTRC 2.0’ (Aug 2017-Aug 2022). £6.7 million awarded to University of Manchester from NIHR.

M. Caulfield
‘NIHR Biomedical Research Centre in Cardiovascular Disease (two CV themes)’ (Apr 2017-Mar 2022) £111 million awarded to London from NIHR (£6.55 million of which was awarded to the Farr Institute).

‘Modelling historical air pollution and dementia/cognitive decline: towards a life course approach’ (Co-I) (Dec 2016-Jan 2018) University of Edinburgh awarded £61,532 from Improving Health with Environmental Data Call NERC/MRC/CSO.

‘Improving Predictive Asthma Algorithms with Modelled Environment Data for Scotland’ (CoI) (Dec 2016-Jan 2018). £63,000 awarded to the University of Edinburgh from Improving Health with Environmental Data Call NERC/MRC/CSO.

A.L. Cox
‘GetAMoveOn: transforming health through enabling mobility’ (Jun 2016- May 2020) £923,685 awarded to London from EPSRC.

S. Cunningham-Burley
‘Guthrie Cards for Research Purposes’ to conduct a citizen’s jury on the use of. Awarded £40,000 from University of Edinburgh/Wellcome Trust.

‘Filling the Void: Public Engagement around a New Model for Access to Research Resources’. CIHR grant worth $530,000, PI Kim McGrail, UBC, Canada.

J. Deanfield
‘Cardio-Oncology Project’ (2017-2022). £1.3 million awarded to London from BHF/CRUK.

Appendices

Research Study on Revascularisation and Angiography’ (2016). £2,000 awarded to London from University of Leicester.


‘NCAP’ (2017-2020). £4.3 million awarded to London from HQIP.

W. Dixon
‘Manchester Day public engagement funding’ (June 2016). £6,000 awarded to University of Manchester from Manchester City Council.

‘Arthritis Research UK public engagement funding for Manchester Day’ (June 2016). £10,700 awarded to University of Manchester from Arthritis Research UK.

‘Wellcome Trust ISSF’ (June 2016). £4,850 to University of Manchester from Wellcome Trust.

‘Transparent and proportionate governance: A trusted use of health data’ (Jan 2017-Jan 2019). £1.2 million awarded to University of Manchester from Arthritis Research UK.

S. Duffy
‘Pilot trial breast cancer awareness’ £30,000 awarded to London from Wolfson Institute.

A. Edwards
‘Evaluating effectiveness, safety, patient experience and system implications of different models of using GPs in or alongside Emergency Departments (GPs in EDs)’ (Mar 2017-Feb 2020). £752,360 awarded to CIPHER from NIHR HS&DR (£18,527 of which was awarded to the Farr Institute).

G. Evans
‘Providing breast cancer risk information as part of national breast cancer screening programme: building an evidence base on benefits and harms to inform a decision to implement’ (2017-2020). £960,000 awarded to University of Manchester from NIHR PGfAR (£300,000 of which was awarded to the Farr Institute).

A. J. Flewitt

D. Ford
‘Examining the effect of education on health and health on education in children and adolescence’ £191,670 awarded to CIPHER from ESRC.

‘Dynamic Insights for Health and Care’ (Apr 16-Mar 18). £57,006 awarded to CIPHER from NERC.

‘UK MS Register’ (2017-2018), £309,267 awarded to CIPHER from MS Society.

‘Full Evaluation of the Supporting People Programme’ (2016-2018) £268,000 awarded to CIPHER from Welsh Government.

E. Ford
‘ASTRODEM: Using astrophysics to close the “diagnosis gap” for dementia in UK general practice’ (Jul 2016-Jun 2018). £94,008 awarded to CIPHER from Wellcome Trust.

P. Foster
‘Big Data Research in Ophthalmology’ US$ 100,000 awarded to London from Alcon Foundation.

‘PhD Studentship – Dr Anurag Garg’ £98,785 awarded to London from Moorfields Eye Charity.

‘Enhanced, Repeat Imaging of UK Biobank Participants’ £35,000 awarded to London from International Glaucoma Association.

‘Biological Pathway Discovery for glaucoma through analysis of genetic and real world data’ £108,336 awarded to London from Moorfields Eye Charity.

‘How do mutations in ubiquitous protein, SPATA 13, lead to Primary Angle-closure Disease in humans?’ £16,750 awarded to London from Moorfields Eye Charity.

‘Automated retinal microvascular quantification as a predictor of cardiovascular disease risk in UK Biobank’ £194,978 awarded to London from BHF.

J. Frank
‘Development Epidemiology: Identifying evidence-based interventions for improving population health and promoting health equity’ (Jul 2016-Jun 2022). University of Edinburgh were awarded $CDN 1.3 million from Canadian Institutes of Health Research.

R. Gilbert

‘Understanding the health needs of mothers and children involved in family court cases’ (Feb 2017-Jan 2020) £394,802 awarded to London from Nuffield Foundation.

‘Head and Heart Study’ (Jan 2017-Jun 2018) £99,463 awarded to London from GOSH.

H. Hemingway
‘Innovative Medicines Initiatives (IMI) Horizon 2020 Big Data for Better Outcomes’ (Mar 2017-Feb 2022) £20 million awarded to London from EU (£1.1 million of which was awarded to the Farr Institute).

‘NIHR Biomedical Research Centres (BRC) Healthcare Informatics, Genomics/Omics, Data Science theme’ (Apr 2017-Mar 2022) £111 million awarded to London from NIHR (£2.9 million of which was awarded to the Farr Institute).

‘Data science Respiratory Innovation Hub’ (Jan 2017-Jan 2022) £1 million awarded to London from GlaxoSmithKline.

‘NHS Hospital Genomics’ (Jan 2017- Jan 2022), £1 million awarded to London from GlaxoSmithKline.

‘NIHR Senior Investigator’ (Apr 2017-Mar 2022). £75,000 awarded to London from NIHR.

A. Hingorani
‘NIHR Senior Investigator’ (Apr 2016-Mar 2021). £75,000 awarded to London from NIHR.
Appendices

C. Hopkins
‘Defining best Management for Adults with Chronic RhinOsinusitis: the MACRO Programme’ £3.2 million awarded to London from NIHR.

‘NIHR programme grant for research into chronic rhinitis’ (2016-2022) £3.2 million awarded to London from NIHR.

K. Jones
‘Mobile Operator-Related Phone & Health Data (MORPHeD) Governance: A review of the information governance implications of combining mobile phone network data with health data’ (Apr 2016-Sep 2017). £45,000 awarded to CIPHER from NERC.

R. Lyons

‘Health outcomes of children and adults with learning difficulties’ (May 2016-Apr 2017). £30,000 awarded to CIPHER/Farr Institute Scotland from MRC.

‘Senior Faculty Award’ (Mar 2016-Mar 2017). £15,000 awarded to CIPHER from Health and Care Research Wales.

‘Evaluation of the Living Well Living Longer programme’ (Mar 2016-Dec 2018) £41,600 awarded to CIPHER from Aneurin Bevan University Health Board.

‘Health Attainment of Pupils in Primary Education Network (HAPPen)’ (Oct 2016-Sep 2019). £57,501 awarded to CIPHER from ESRC.

‘Creating a profile of population-based stage at diagnosis by cancer type and geography in Wales, and one-year relative by stage for the main cancer types’ (Mar 2017-Feb 2018). £53,905 awarded to CIPHER from Cancer Research.

‘Informatics work to underpin the ARCH Regional Collaboration for Health’ (Nov 2016-Feb 2017). £40,000 awarded to CIPHER from ARCH & Swansea University.

R. Lyons & S-M. Zhou

G. Nenadic
‘ISACC’ (Feb 2017-Sep 2017). £27,000 awarded to University of Manchester from the Industry: IMO.

T. Paljarvi
Feasibility of using consumer-targeted activity monitors in collecting physical activity data within the HealthWise Wales cohort: A pragmatic pilot study’ (Mar 2016-Dec 2016). £22,731 awarded to CIPHER from Wellcome Trust.

S. Paranjothy
‘Morbidity following urinary tract infection in young children: an electronic record-linked cohort study’ (Oct 2016-Sep 2018). £248,000 awarded to CIPHER from Health and Care Research Wales (£244,111 of which was awarded to the Farr Institute).

‘Evaluation of the outcomes of the family nurse partnership programme in Scotland’ (Jan 2016-Sep 2018). £183,000 awarded to CIPHER from Scottish Government.

N. Peek
‘Wearable clinic’ (Mar 2017-Feb 2017). £2 million awarded to University of Manchester from the EPSRC.

‘CHAMP Web-based Support Tools’ (Dec 2016-Apr 2017). £7,900 awarded to University of Manchester from City in the community (charity arm of MCFC).

S. Petersen

‘Genome-wide association scanning of cardiovascular functional traits in UK Biobank’ (Apr 2017-Mar 2021) £203,720 awarded to London from The Medical College of Saint Bartholomew’s Hospital Trust.

Ceri Phillips
‘Cost of not treating Multiple Sclerosis with Disease modifying therapy’ (Mar 2017-Feb 2018). £149,929 awarded to CIPHER from the MS Society (£26,884 of which was awarded to the Farr Institute).

K. Pickett
‘MRC Research Grant’ (Sep 2016-Aug 2021). £3.6 million awarded to University of Manchester from MRC-UK.

K. Pritchard-Jones
‘Support for establishing a Centre for Cancer Outcomes as part of the UCLH Cancer Collaborative (national cancer vanguard new care models programme)’ (2016-2017). £122,500 awarded to London from UCLH Charity.

A. Radford
‘SAVSNET scaleup: Infrastructure’ (Apr 2017-Mar 2020). £700,000 awarded to University of Manchester from BBSRC.

A. Renehan
‘Obesity and Cancer TOgether imPact Upon Survival (OCTOPUS) consortium ‘Cancer e-lab’: a federation meta-analysis of trial data’ (Feb 2017-Feb 2019). £500,000 awarded to University of Manchester from World Cancer Research Fund.

J. Robson
‘Low birthweight’ £2,000 awarded to London from University of East London.

‘Asthma predictive score’ £13,000 awarded to London from CLARHC.

‘Clinical Effectiveness Group’ £850,000 awarded to London from Newham, Tower Hamlets, City and Hackney CCG and Public health.

S. Rodgers
‘Air pollution, pollen, asthma and hayfever and their interactive effect on cognitive development: an environment and health data linkage feasibility study’ (Nov 2016-Oct 2017). £109,851 awarded to CIPHER from NERC.

J. Sampson
‘Whole Genome Sequencing for Health and Wealth: A bid to establish the Wales Genomic Medicine Centre to contribute to the UK 100,000 Genomes Project and Strengthen Genomics in Wales. (Kerina Jones, Co-I)’ (Dec 2016-Nov 2018). £45,133 awarded to CIPHER from the MRC.
M. Sharpe (PI, Oxford University)
‘Pragmatic randomised trial to compare the effectiveness and cost-effectiveness of Proactive Liaison Psychiatry with usual care’ (2016-2018). £1.9 million awarded to University of Manchester (£20,000 of which was awarded to the Farr Institute) from NIHR HS&DR.

L. Smeeth
‘Impact of the media interest in side effects on the initiation and cessation of statins in the UK’ (Apr 2016-Mar 2017). £55,000 awarded to London from BHF.

‘Glycaemia and chronic disease: harnessing UK Biobank and e-Health linkage to quantify risks, explore mechanisms and determine treatment impacts’ (Sep 2016-Aug 2021). £698,136 awarded to London from Diabetes UK & British Heart Foundation.

‘Role of persisting neurotropic viruses in neurological diseases of ageing’ (Oct 2016-Sep 2020). £704,627 awarded to London from Wellcome Trust.


‘Real world effects of medications for chronic obstructive pulmonary disease’ (Sep 2016-Aug 2019). £350,000 awarded to London from NIHR.


‘Chronic Kidney Disease of unknown cause (CKDu) in disadvantaged communities in low-and-middle income countries (LMICs)’ (Apr 2017-Mar 2019). £566,066 awarded to London from MRC.

‘Using linked electronic health data to improve eczema diagnosis and outcomes’ (Mar 2017-Feb 2022). £1.3 million awarded to London from Wellcome Trust.

‘Preserving Antibiotics through Safe Stewardship: PASS’ (Apr 2017-Mar 2020). £1.6 million awarded to London from ESRC.

M. Sperrin
‘CHAMP weighing and measuring evaluation’ (May 2016-August 2017). £60,000 awarded to University of Manchester from the City in the community (charity arm of MCFC).

A. Timmis
‘Cardiac care pathways’ £124,000 awarded to London from UCLPartners.

T. van Staa
‘Building Rapid Interventions to reduce microbial resistTance and over-prescribing of antibiotics (BRIT)’ (2016-2018). £700,000 awarded to University of Manchester from Department of Health (Connected Health Cities)

Gern; Tachinardi (PI, University of Wisconsin)
‘CREW project’ (Sep 2016-Aug 23). $100 million awarded to University of Manchester ($1 million of which was awarded to the Farr Institute) from NIH, USA.

C. Taylor
‘NewMind Plus Network’(2016-2019). £630,000 awarded to University of Manchester from EPSRC.

C. Taylor and J. Ainsworth
‘CityVerve’ (2016-2018). £10 million awarded to University of Manchester from IUK.

P. Taylor
‘Vital@Home’ (Jan 2016-Dec 2016) €184,483 awarded to London from European Institute of Technology.

‘UK EMR Medical Retina collaboration’ (2016-2019). $1.9 million awarded to London from European Novartis.

M. Turner

S. van der Veer
‘Improving the reporting of uraemic pruritus in dialysis’ (Aug 2016-Feb 2018). £39.5k awarded to University of Manchester from Kidney Research UK

P. Whelan and J. Ainsworth
‘ACTIF II’ (2016—2018). £177,000 awarded to University of Manchester from a Suite of funders.

James Wildman
‘CF HealthHub part II’ (Sep 2016-Feb 2020). £1.6 million awarded to University of Manchester from the NIHR.

E. Williamson
‘Real World Effects of Medications for Chronic Obstructive Pulmonary Disease’ (2016-2019) £291,953 awarded to London from NIHR.

I. White
‘Imputation by chained equations for data that are MNAR’ (2016-2018). £165,000 awarded to London from MRC.

X. Xie and S-M. Zhou
‘Unravelling polypharmacy: Mining the complex interaction patterns between medications for enhanced patient care’ (Oct 2016-Sep 2019). £56,000 awarded to CIPHER from Swansea University (£28,000 of which was awarded to the Farr Institute).
Appendix 4: Patient and Public Involvement and Engagement activities

For Researchers

Public Trust and Engagement in Research Course, 26 April 2016, Farr London
13th World Congress of the International Association of Bioethics, 14-17 June 2016, Edinburgh
MQ Mental Health Data Science Group Meeting, 12 September 2016, London
Scottish Mental Health Research Network Annual Scientific Meeting, 4 October 2016, Edinburgh
Innovate UK Conference, 2-3 November 2016, Manchester
BSA Medical Sociology Scotland Group Meeting, 24 November 2016, Dundee
Public Academic Research Colloquium, 29-30 November 2016, Washington DC
Drop-in Advice Sessions, ongoing, Edinburgh
Drawing on data to transform lives: improving services for vulnerable adolescents, 1 Feb 2017, London

For the Publics

Using Health Record Data – from Discovery to Public Health, UCL Festival for Digital Health, 1 March 2016, London
Citizens Jury’ Feedback Event, 10 March 2016, Manchester
Public Health England Hackathon, 18-19 May 2016, Manchester
The Times Cheltenham Science Festival, 7–12 June 2016, Cheltenham
GCHQ Cyberacademy (Schools engagement), 10 June 2016, Edinburgh
Patient Portals: Co-designing the Future Workshop, 18 June 2016, Manchester
Research Open Day for Patients and the Public, 28 June 2016, UCL/UCLH
Barts & Queen Mary Science Festival, 6 July 2016, London
Public debate: Who owns my genome?, 7 July 2016, Edinburgh
in2science Work Placement Scheme (Schools engagement), 11-15 July 2016, London

Bluedot Festival, 22-24 July 2016, Cheshire
#BritainBreathing, Pint of Science, 26 July 2016, Manchester
Edinburgh Fringe Festival, 12 and 19 August 2016, Edinburgh
International Population Data Linkage Network Conference, 22-26 August 2016, Swansea
British Science Festival, 7-9 September 2016, Swansea
CHAMP Trumps Game, 12 September 2016, Manchester
Carmarthenshire 50+ Forum Annual Event, 16 September 2016, Carmarthenshire
Technique After Hours, 29 September 2016, Cardiff
Body Beats Hackathon, 1 October 2016, Manchester
Health and Care Research Wales Conference, 13 October 2016, Cardiff
Collaborate, 19 October 2016, Swansea
Manchester Science Festival, 23 October 2016, Manchester
Science Museum Lates: Contagion, 26 October 2016, London
Community Reporter Training, October/November 2016, Manchester
ACTIVE Project Fitness Testing and Baseline Testing, 17 and 30 November 2016, Swansea
Famelab Heat, 8 December 2016, Edinburgh
ACTIVE Project Peer Mentoring Social Media Workshop, 9 December 2016, Swansea
Drawing on Data to Transform Lives: Improving Services for Vulnerable Adolescents, 1 Feb 2017, London
ACTIVE Project Launch, 06 January 2017
The Involving People Network Evaluation Event, 11 January 2017
Swansea University District Nursing Students visit, 27 February 2017