

Making game-changing improvements in the health of patients and populations

Ronan Lyons, Professor of Public Health, Swansea University

SA-NT Datalink, 3rd December 2018

r.a.lyons@Swansea.ac.uk



A new national Institute for health data science



Launched in April 2018, our UK team of experts develop and apply cutting-edge data science approaches to clinical, biological, genomic and other multi-dimensional health data to address the most pressing health research challenges facing the public.



We are funded by the Medical Research Council, the British Heart Foundation, the National Institute for Health Research, the Economic and Social Research Council, the Engineering and Physical Sciences Research Council, Health and Care Research Wales, Health and Social Care Research and Development Division (Public Health Agency, Northern Ireland), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, and Wellcome.



About me

Dubliner, living in Swansea

“We are all in the gutter but some of us are looking at the stars”

- Background in hospital and emergency medicine, epidemiology
- Data interests from injury surveillance, 1993-
- Involved in injury surveillance in UK, Europe and internationally (Chair Injury ICE for US NCHS/CDC)
- Led to broader involvement in health informatics
 - Health Informatics Research Unit, 2007-
 - SAIL Databank, 2009-
 - **MRC** centres: CIPHER, DECIPHer, DPUK, Stratified Medicine, 2012-, HDRUK 2018-
 - **ESRC**: ADRC-Wales, cohort linkage 2013-
- Member of MRC MRP, Wellcome Science Interview Panel, BHF Data Science Implementation Group, OSCHR health informatics...



Increasingly, I know less and less about more and more

HDR UK triple aim

Scientific programmes

We will integrate data science with biomedical and health science expertise to perform ground-breaking research, with an initial focus on data analytics, precision medicine, 21st century clinical trials and modernising public health.

Training the next generation

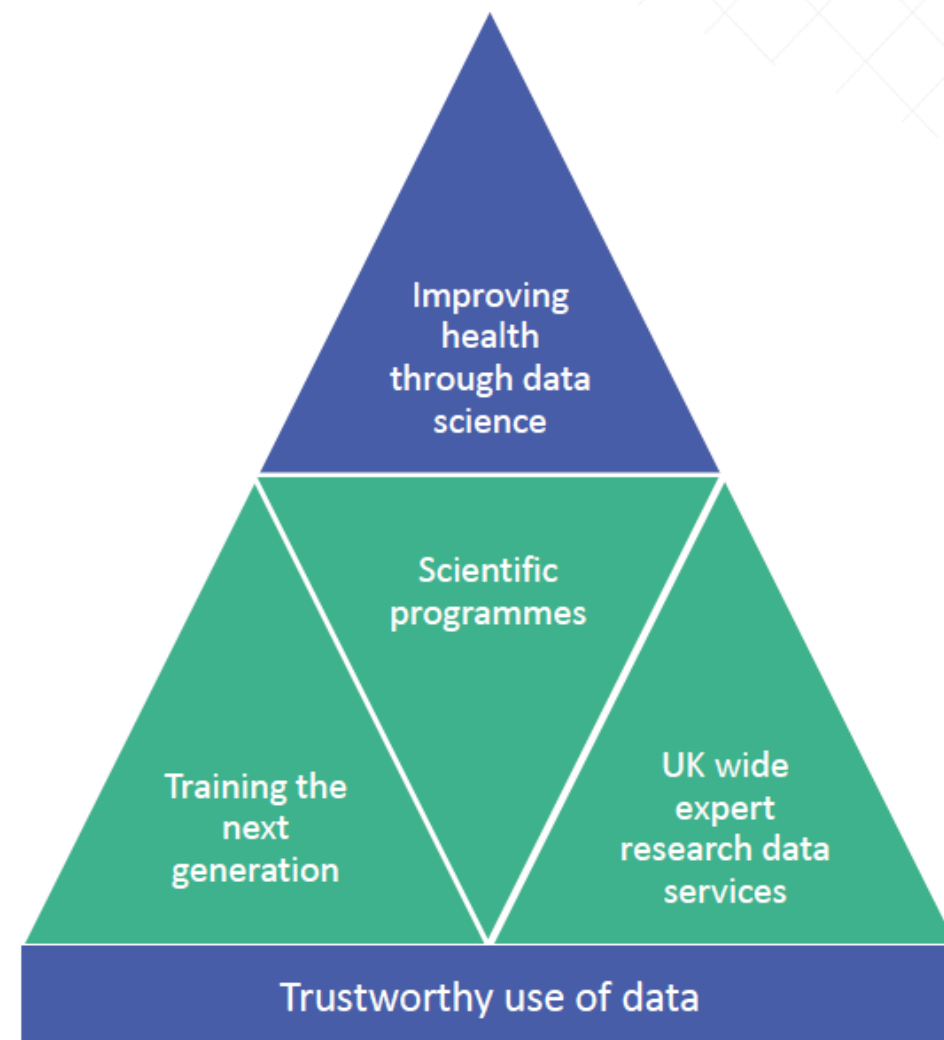
Our training programme will embrace novel approaches to research training and mentorship to foster a cadre of health data science researchers, on a substantial scale.

UK wide expert research data services

We will develop cutting-edge technologies and trusted research platforms that acquire, store, represent, and process large, multi-dimensional research data.

Trustworthy use of data

We will work in partnership with the public, funders, social scientists and legal/ethnicity experts to champion the trustworthy use of data.



Why now? Complex environment

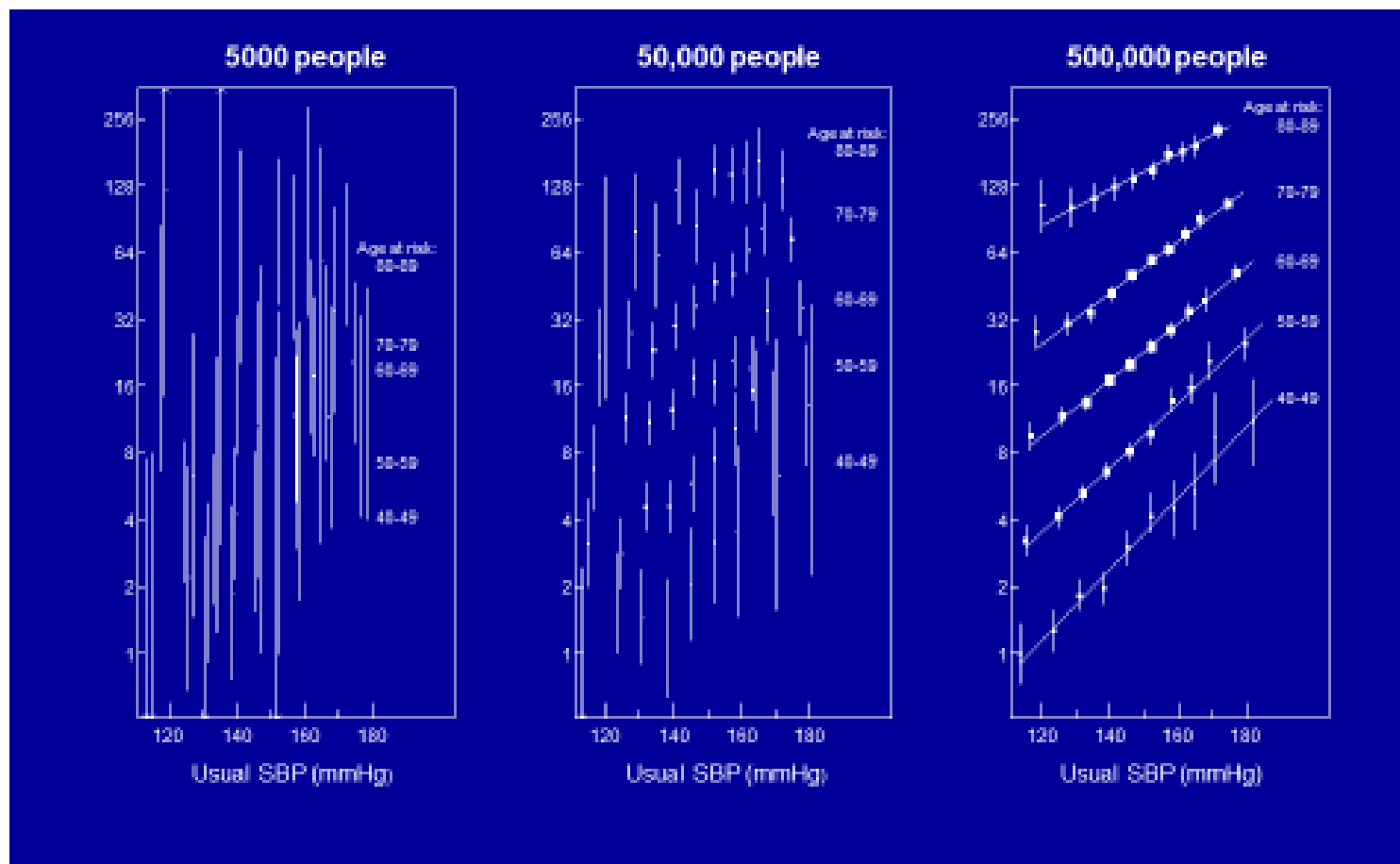
Establishing processes for large-scale health outcome phenotyping



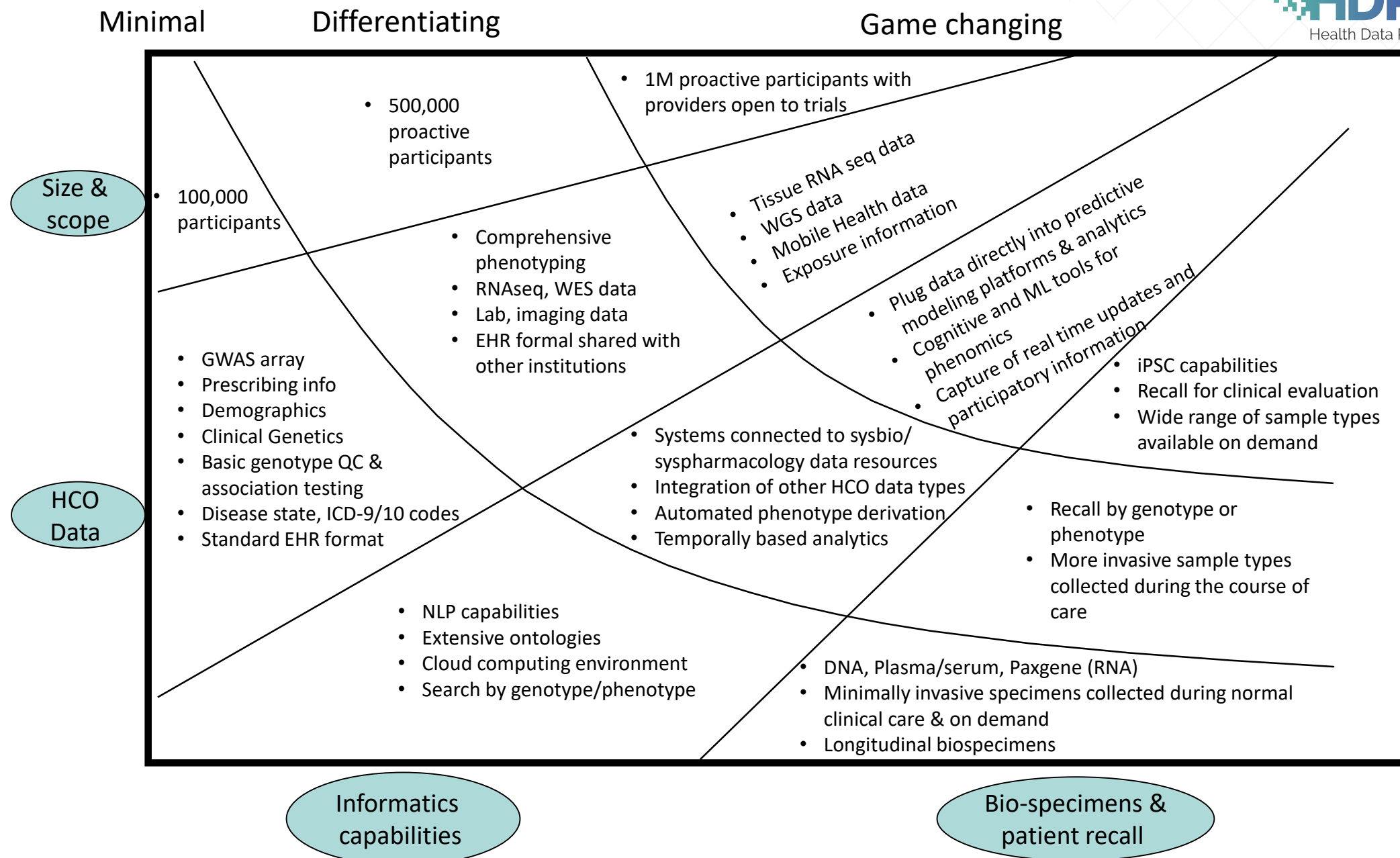
Interoperability: to work across systems with no additional effort

Enormous scale required to establish clinical utility

Prospective studies need to be large: heart disease mortality vs systolic blood pressure (data from 1M people with 12 year follow-up) – Rory Collins



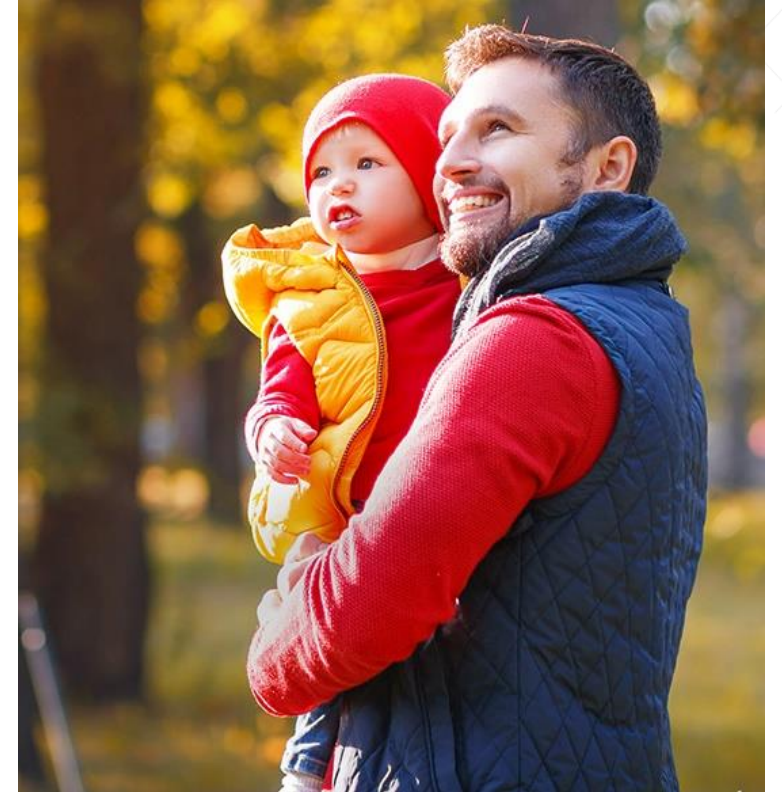
Emerging Requirements for Genomic Resources



Desired impact

Our mission is to make game-changing improvements in the health of patients and populations through data science research and innovation.

- Discovery science
- Prevention
- Earlier diagnosis
- Safety and efficacy of therapeutics
- Driving clinical research
- Precision medicine



An outstanding partnership

Our initial investment supports six Sites. Each has world-class expertise; a track record in using health data to derive new knowledge, scientific discovery and insight; and works in close partnership with NHS bodies, industry and the public to translate research findings into benefits for patients and populations.



1. Wales and Northern Ireland (Swansea and Queen's University Belfast)

2. Midlands (Birmingham, Leicester, Nottingham, Warwick)

3. Scotland (Glasgow, Edinburgh, Dundee, Aberdeen, Strathclyde, St Andrews)

4. London (Imperial, Kings, London School of Hygiene and Tropical Medicine, Queen Mary, UCL)

5. Oxford

6. Cambridge (EBI, Sanger, Cambridge University)

HDR UK Research Directors



- National and regional responsibility
- Senior Scientific Leadership Committee

Simon Ball
Midlands



John Danesh
Cambridge



Harry Hemingway
London



Martin Landray
Oxford



Ronan Lyons
Wales/NI

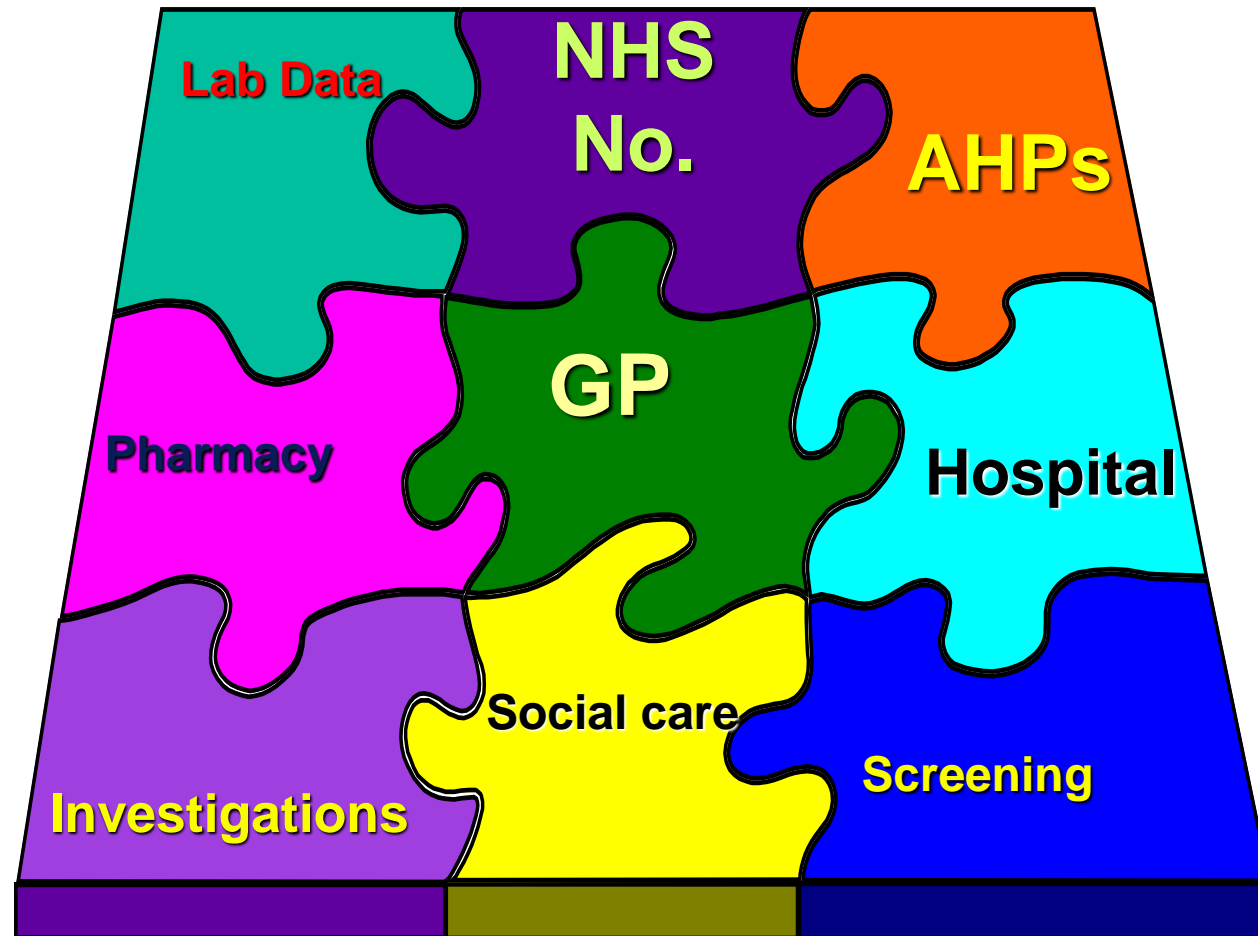


Cathie Sudlow
Scotland

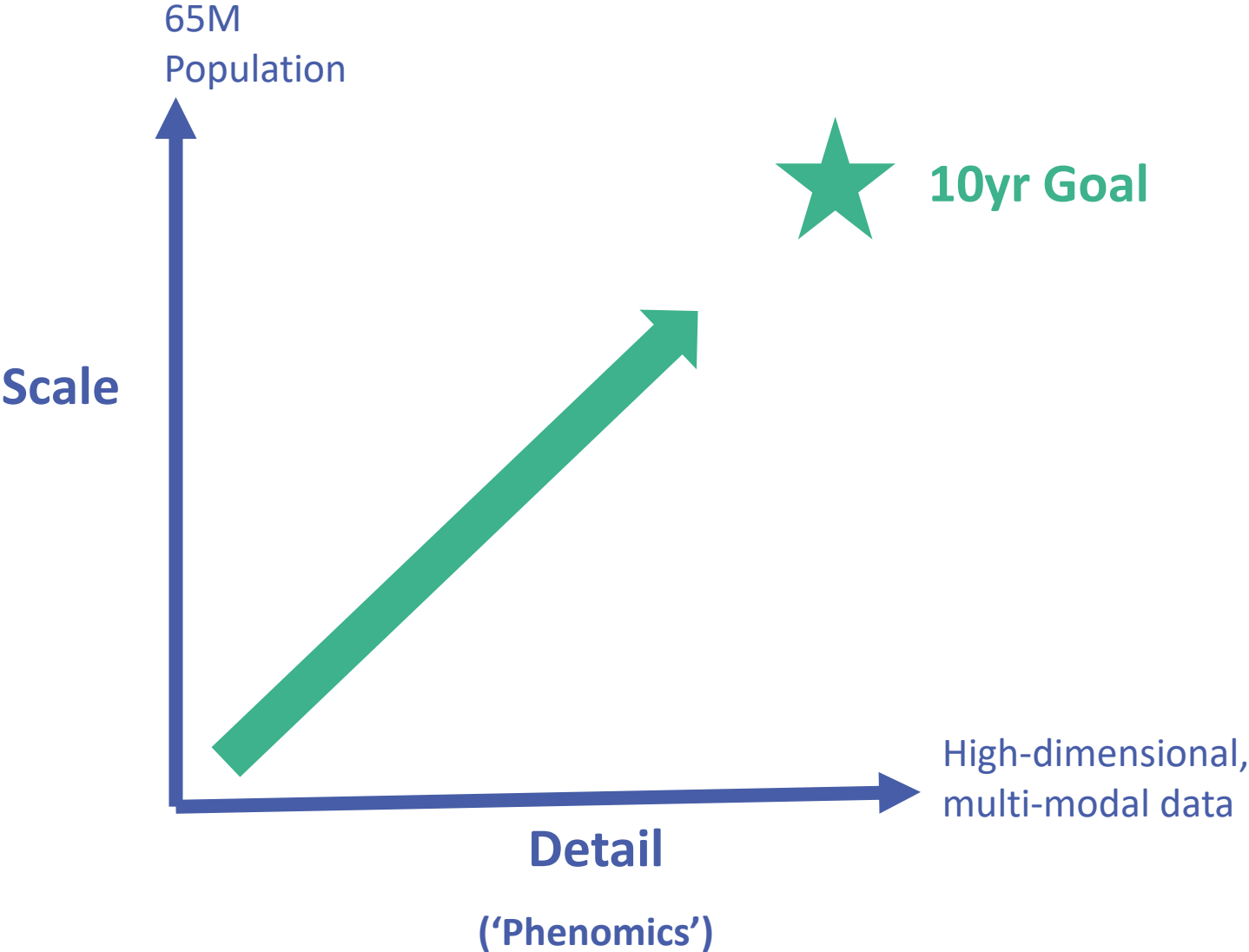


Linking Data

key to seamless care and research



Our long-term goal



Scientific priorities

Actionable Health Data Analytics:

- Structured and unstructured (e.g. imaging, text) data for derivation of new or deep phenotypes
- Adding value at scale to existing world-leading cohorts in the UK – 2.5M
- Demonstrating system-wide opportunities for research that improves quality of care

Precision Medicine:

- Enable large scale, high-throughput research combining omic data with EHRs: Integromics

21st Century Trial Design:

- Transform Phase II – Phase IV clinical trials including ‘real world evidence’ studies

Modernising Public Health: towards prevention and early intervention

- Population e-cohorts: link health and administrative datasets across multiple environments
- New technologies, from sensors to wearable devices to artificial intelligence

The Alan Turing Institute

- **Purpose:** To provide leadership at the computer science and health data science interface
- **Timeline:** Joint appointment of Chris Holmes with ATI and HDR UK commenced April 2018
- **Status:** Partnership in place - Sir Adrian Smith
- **Focus:** Training, Grand challenges

The Alan Turing Institute



Pioneers in training the next generation of data science leaders, shaping the public conversation, and pushing the boundaries of data science

Developing the research landscape

- Directors agreed plan for collaborative science at scale
- Call for *Implementation Projects*
- Iterative development of cross site proposals to December 2018
- NLP derived phenotypes
- Multi-morbidity
- Multi-omic cohorts for discovery science
- Board approval required: £3.6M initial budget



Training future leaders in health data science

Our training strategy:

- Develop flexible training opportunities
- Address unmet skills needs
- Establish clear career structures
- Develop appropriate reward and recognition frameworks
- Enable career progression through intellectual leadership and independence



UK wide expert research data services

- Guiding principles for HDR UK national health data research infrastructure
- MRC £10M capital opportunity
- Blueprint of national strategy by April 2019
- Building on SAIL, Scottish system, Local Health and Care Record Exemplars (LHCREs)



Wales and Northern Ireland - Embracing Partnership in the Health Data Revolution

Ronan Lyons, University of Swansea
HDR Wales-Northern Ireland Director

Mark Lawler, Queen's University Belfast
HDR Wales-Northern Ireland Associate Director

Wales and Northern Ireland - aligned to HDR UK priorities

Actionable Analytics

Core underpinning activities

Enhance our platforms for distributed team science:

SAIL, UKSeRP, Self2Mind, GA4GH

Add value to UK cohorts:

DPUK, UKB, ELGH, ALSPAC

Conduct system-wide patient journeys and analytics

Create even deeper phenotypes:

Clinical correspondence (NLP)

Lab results

Images

Precision Medicine

Genomics/bioinformatics

Refractory asthma

Colorectal cancer

Biomarker discovery and validation

Digital pathology

Centres for Nano Health (CNH) and Nano-Medicine

Modernising Public Health

Co-develop and combine population cohorts from Wales/NI with other areas of UK

Observational studies

Upstream and downstream *population level* inequalities research

Evaluations of natural experiments

Trials

CVD

Cancer

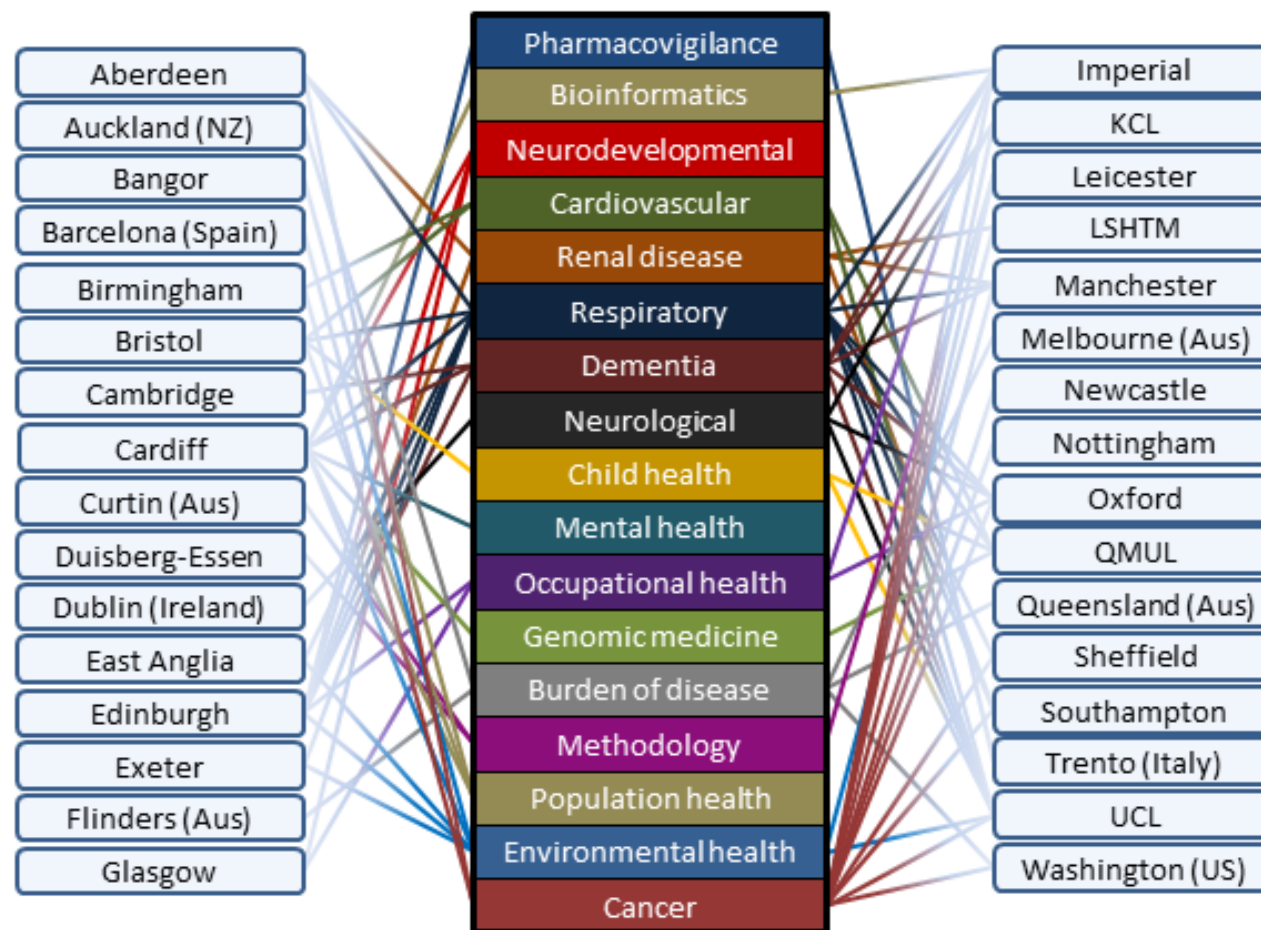
Asthma

Dementia

Injury

Multi-morbidity

Collaborative



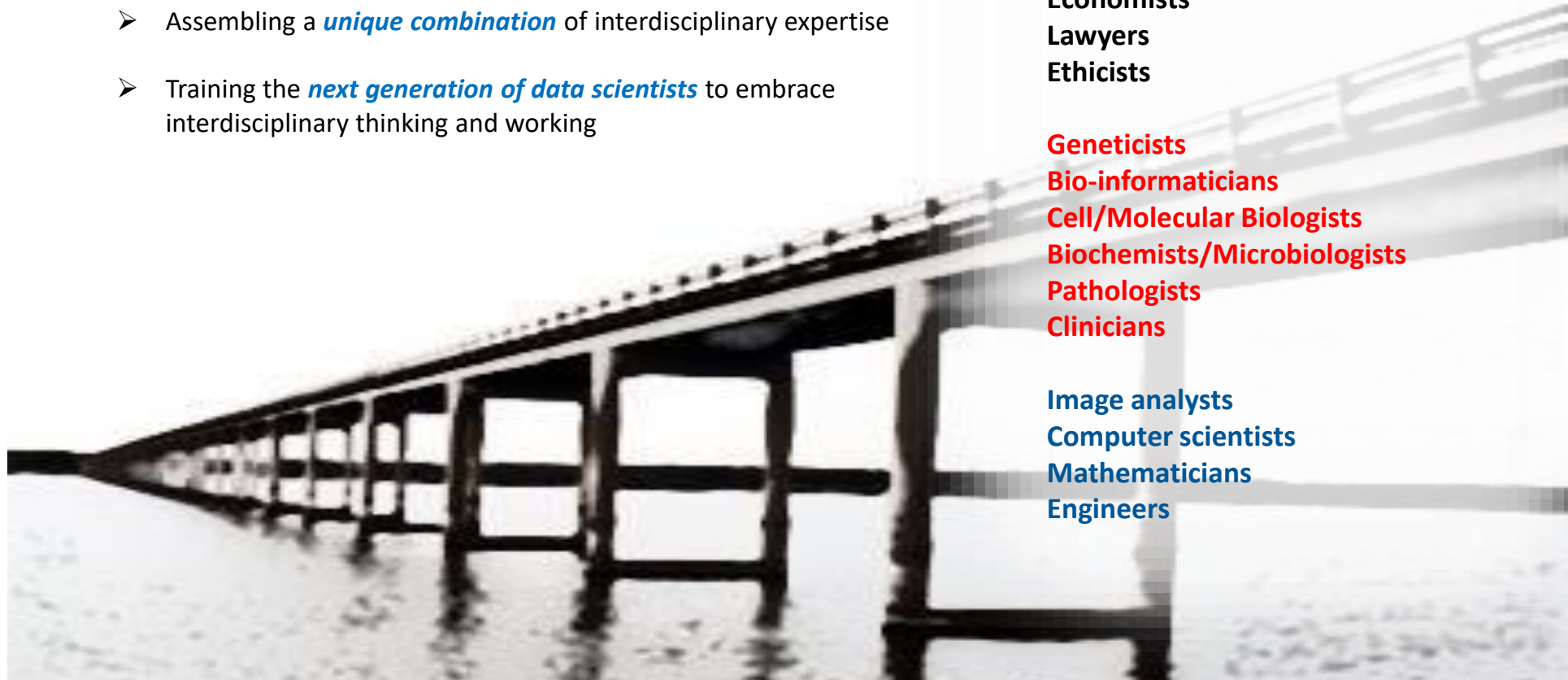
Transforming delivery Transcending disciplines

- A *Step-change* in **SCALE** and **PACE** to tackle **GLOBAL CHALLENGES**
- Bringing new disciplines *into the HDR domain*
- Assembling a *unique combination* of interdisciplinary expertise
- Training the *next generation of data scientists* to embrace interdisciplinary thinking and working

Epidemiologists
Statisticians
Health service researchers
Clinicians
Geographers
Demographers
Social scientists
Economists
Lawyers
Ethicists

Geneticists
Bio-informaticians
Cell/Molecular Biologists
Biochemists/Microbiologists
Pathologists
Clinicians

Image analysts
Computer scientists
Mathematicians
Engineers



Fellowship Programme



- UKRI Fellows
- **Jason Carson** (Nitiarasu). Non-invasive assessment and management of coronary heart disease
- **Jonathon Kennedy** (Brophy). Predicting the diagnosis of inflammatory and autoimmune arthritis
- **Matthew Alderdice** (Lawler). Facilitating data integration to empower colorectal cancer discovery research
- **David Wright** (O' Reilly). Data driven public health approaches for diabetic retinopathy and age-related macular degeneration



‘Safe Haven’ for pseudonymised data (Welsh population)

A world leading, privacy protecting research environment

Globally accessible*

Linked, longitudinal de-identified data on Welsh population
5 million people – 20 billion rows of data

Over 20 core national datasets

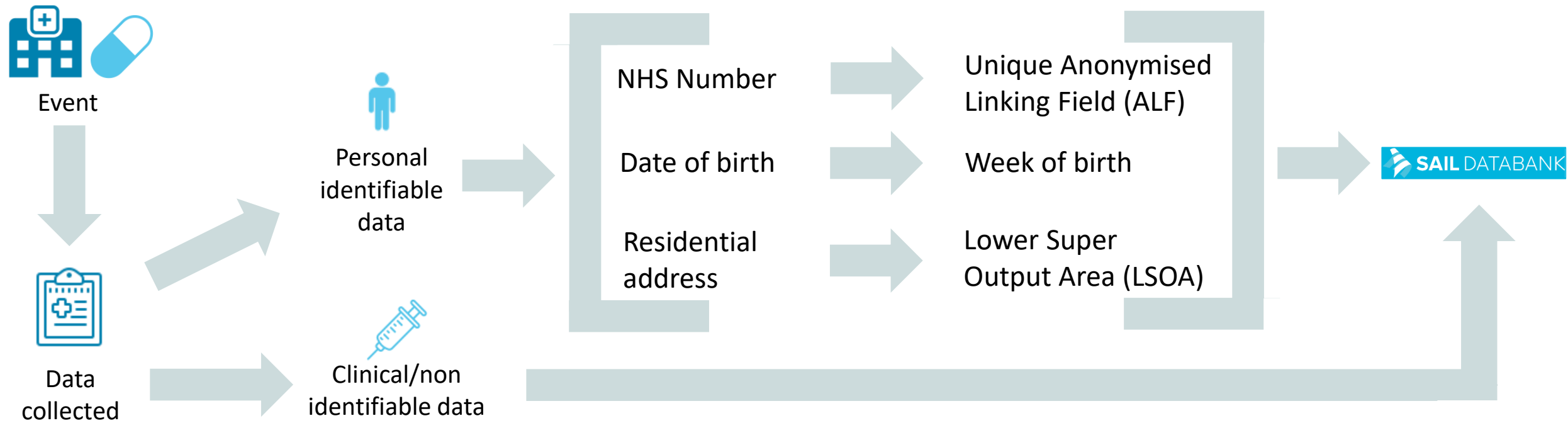
80% GPs provide data

200+ project specific datasets

Non health service and environmental data

Project defined datasets/specific encryption

Remote analysis/manual review of results



SAIL DATABANK

Demographic Data

- Age
- Gender
- Residency
- Socio-economic status

Health Data

- General Practice electronic records
- Emergency Department attendances
- Hospital admissions including critical care received
- Outpatient appointments

Education Data

- Key Stage 4 examination scores
- Special Education needs
- Attendance
- Free school meal eligibility
- Percentage of free school meal uptake
- Number of job vacancies

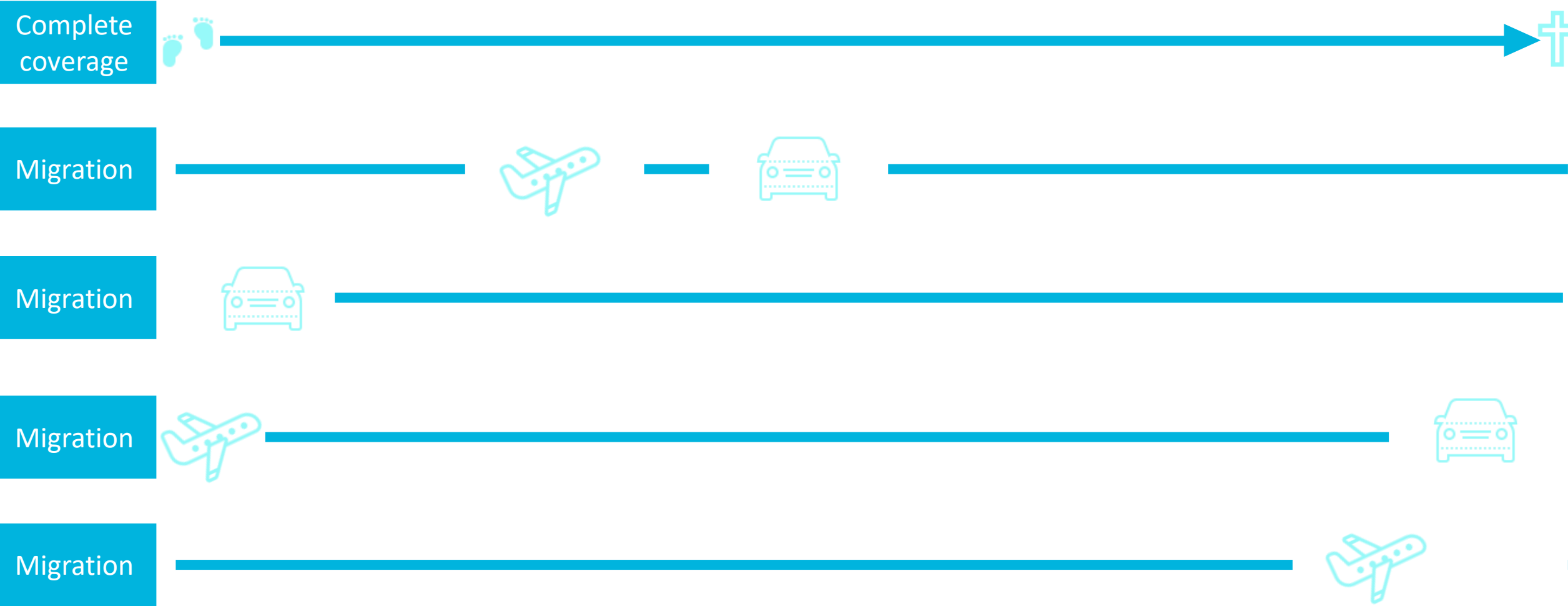
Pollution Data

Pollen Data

Lifecourses in SAIL

Welsh
Demographic
Service
(WDS):

Manages demographic data (name, address, date of birth) of Welsh residents who are registered with a General Practice in Wales.



Population and clinical e-Cohorts

THE LANCET Public Health

Volume 3, Issue 6, June 2018, Pages e279-e288

[open access](#)

Articles

Risk of emergency hospital admission in children associated with mental disorders and alcohol misuse in the household: an electronic birth cohort study

Prof Shantini Paranjthy PhD ^{a, b, c, d, e}, Annette Evans MSc ^f, Amrita Bandyopadhyay BTEch ^g, Prof David Fone MD ^{a, b, c, d}, Behnaz Schofield PhD ^h, Prof Ann John MD ^{b, c}, Prof Mark A Bellis DSc ^g, Prof Ronan A Lyons MD ^{b, c}, Daniel Farewell PhD ^{a, c}, Sara Jayne Long PhD ^g

PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Gestational Age, Birth Weight, and Risk of Respiratory Hospital Admission in Childhood

Shantini Paranjthy, Frank Dunstan, William J. Watkins, Melanie Hyatt, Joanne C. Demmler, Ronan A. Lyons and David Fone
Pediatrics 2013;132:e1562; originally published online November 18, 2013;
DOI: 10.1542/peds.2013-1737

Open access

Research

BMJ Open Acute kidney injury in the UK: a replication cohort study of the variation across three regional populations

Simon Sawhney,^{1,2} Heather A Robinson,^{2,3} Sabine N van der Veer,^{2,3} Hilda O Hounkpatin,^{2,4} Timothy M Scale,^{2,5} James A Chess,^{2,5} Niels Peek,^{2,3} Angharad Marks,^{1,2} Gareth Ivor Davies,^{2,5} Paolo Fraccaro,^{2,3} Matthew J Johnson,^{2,4} Ronan A Lyons,^{2,5} Dorothea Nitsch,^{2,6} Paul J Roderick,^{2,7} Nynke Halbesma,^{2,8} Eve Miller-Hodges,^{2,9} Corinda Black,^{1,2} Simon Fraser^{1,2}

Epilepsia®

Official Journal of the International League Against Epilepsy

Full-Length Original Research | [Free Access](#)

Epilepsy and deprivation, a data linkage study

William O. Pickrell , Arron S. Lacey, Owen G. Bodger, Joanne C. Demmler, Rhys H. Thomas, Ronan A. Lyons, Phil E. M. Smith, Mark I. Rees, Mike P. Kerr

First published: 02 March 2015 | <https://doi.org/10.1111/epi.12942> | Cited by

icott@swansea.ac.uk



Schizophrenia Research

Volume 166, Issues 1–3, August 2015, Pages 131–

136

A national population-based e-cohort of people with psychosis (PsyCymru) linking prospectively ascertained phenotypically rich and genetic data to routinely collected records: Overview, recruitment and linkage

Keith Lloyd ^{a, b, c, d, e}, Joanna McGregor ^{a, b}, Ann John ^{a, b}, Nick Craddock ^c, James T. Walters ^c, David Linden ^c, Ian Jones ^c, Richard Bentall ^d, Ronan A. Lyons ^{a, b}, David M. Evans ^{a, b}, Michael J. Owen ^c

6

Respiratory epidemiology

BMJ Open
Respiratory
Research

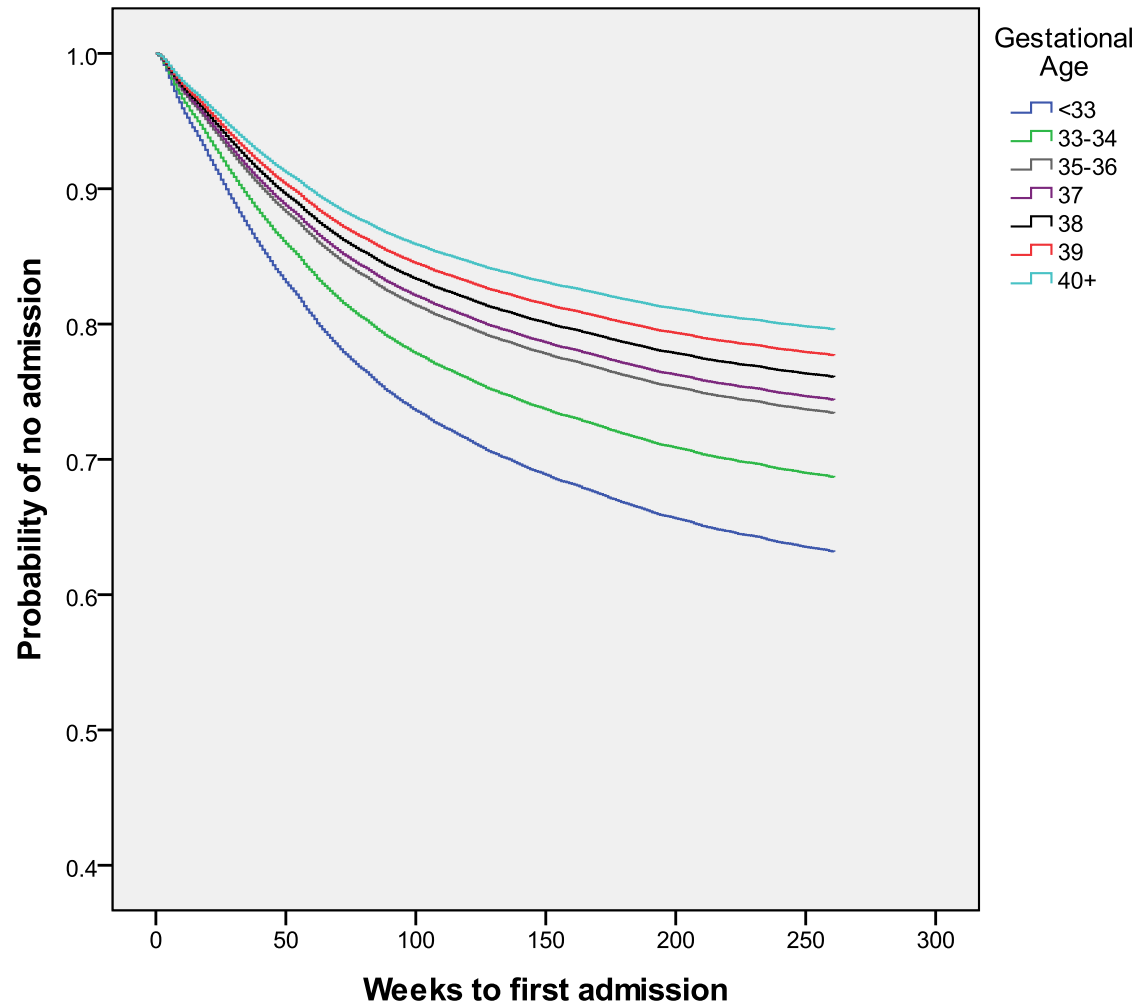
Childhood asthma prevalence: cross-sectional record linkage study comparing parent-reported wheeze with general practitioner-recorded asthma diagnoses from primary care electronic health records in Wales

Lucy J Griffiths,¹ Ronan A Lyons,² Amrita Bandyopadhyay,² Karen S Tingay,² Suzanne Walton,² Mario Cortina-Borja,² Ashley Akdeniz,² Helen Bedford,² Carol Decoteaux^{1,2}

274328-Childhood asthma prevalence: cross-sectional record linkage study comparing parent-reported wheeze with general practitioner-recorded asthma diagnoses from primary care electronic health records in Wales

Wales Electronic Cohort for Children

Time to the first emergency respiratory hospital admission by gestation



- Risk decreased with each successive week in gestation up to 40 – 42 weeks.
- Risk further increased for babies that were small for gestational age.
- The increased risk is small for late preterm infants but the number affected is large and will impact on healthcare services.

Wales: strength in evaluation of natural experiments

SAIL's ALF and RALF system with linkage to social and environmental data is unique

Large scale evaluations of:

- Housing improvement
- Energy efficiency
- Alcohol outlets
- School Breakfasts
- Pollution and Health
- Green/Blue space
- Neighborhood regeneration

Use of data linkage to measure the population health effect of non-health-care interventions



Ronan A Lyons, David V Ford, Laurence Moore, Sarah E Rodgers

Health is largely determined by the interactions between social, economic, and physical environments and inherited characteristics, which together affect exposures and behaviours.^{1,2} Despite an improving evidence base showing the cost-effectiveness of disease prevention with non-health-care interventions, most investments into improving health continue to focus on health service treatment of acute and chronic illnesses, with little mainstreaming of effective preventive interventions other than vaccination or screening.³

Even within the prevention model there is a tendency towards investment in medical models of individual behaviour change, despite growing evidence showing that environmentally based interventions often provide

records and the development of methods for the protection of privacy have revolutionised capabilities. A thriving international data linkage network now exists, with particular strengths in Australia, Canada, Scandinavia, and the UK. The Western Australia Data Linkage System has been operational since 1995, and has been at the forefront of developments. It has contributed to a large reduction in the need for identifiable data in research, while supporting an increasing number of studies and showing more than a ten-fold return on research investment.⁴ The long-term follow-up of the West of Scotland Coronary Prevention Study⁵ provides an excellent example of the cost-effectiveness of the data linkage system within the UK. Evidence from that study

Published Online
November 28, 2013
[http://dx.doi.org/10.1016/S0140-6736\(13\)62750-X](http://dx.doi.org/10.1016/S0140-6736(13)62750-X)
Centre for Improvement in Population Health through E-records Research, Swansea University, Swansea, UK
(Prof R A Lyons MD, Prof D V Ford MBA, S E Rodgers PhD) and DECIPHER Research, Cardiff University, Cardiff, UK (Prof L Moore PhD)
Correspondence to: Prof Ronan A Lyons, Swansea

PUBLIC HEALTH RESEARCH

VOLUME 6 ISSUE 5 MARCH 2018
ISSN 2050-4381



The health impacts of energy performance investments in low-income areas: a mixed-methods approach

Wouter Poortinga, Sarah E Rodgers, Ronan A Lyons, Pippa Anderson, Chris Tweed, Charlotte Grey, Shiyi Jiang, Rhodri Johnson, Alan Watkins and Thomas G Winfield

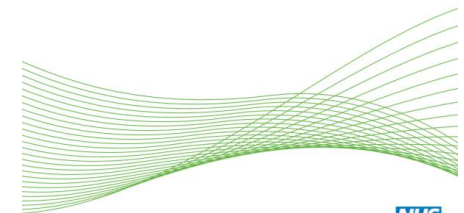
ACCEPTED MANUSCRIPT

Improving Mental Health through the Regeneration of Deprived Neighborhoods: A Natural Experiment

James White ✉, Giles Greene, Daniel Farewell, Frank Dunstan, Sarah Rodgers, Ronan A. Lyons, Ioan Humphreys, Ann John, Chris Webster, Ceri J. Phillips, ... Show more

Am J Epidemiol kwx086. DOI: <https://doi.org/10.1093/aje/kwx086>

Published: 09 May 2017



DOI 10.3310/gh06050

NHS
National Institute for
Health Research



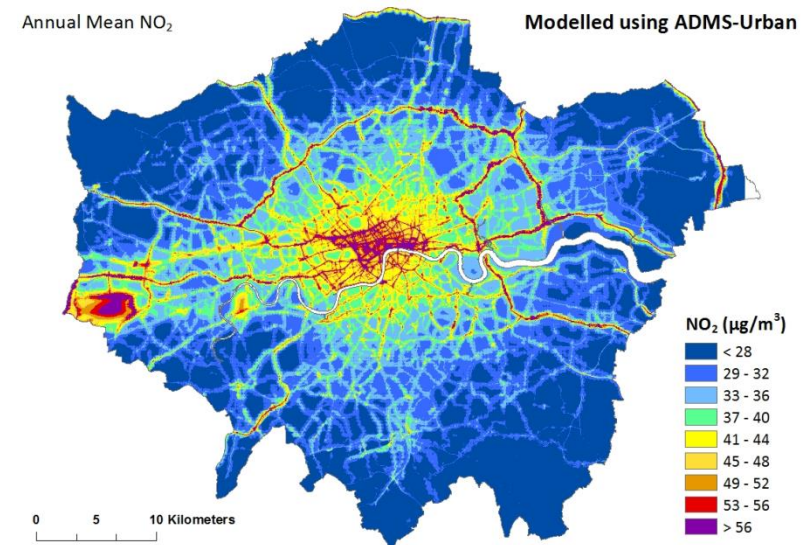
- GIS density calculations:
 - Modelled access by outlet type
 - 10 minute walk
 - 10 minute drive
 - Spatio-temporal interaction gravity model
 - 2.5 billion calculations

Changes in walking distance related to admissions from alcohol related conditions, injury and violent crime



Air pollution, pollen, asthma and hayfever and their interactive effect on cognitive development: an environment and health data linkage feasibility study

- Sarah Rodgers (Swansea University, SU)
- Ruth Doherty (University of Edinburgh)
- Paul Wilkinson and Ai Milojevic (LSHTM)
- Iain Lake (UEA)
- Anna Mavrogianni (UCL)
- Damon Berridge (SU)
- Gwyneth Davies (SU)
- Lorraine Deardon (IFS)
- David Carruthers (CERC)



Chief Scientist Office of Scotland, the Medical Research Council and the Natural Environment Research Council. Grant Ref: R8/H12/83/NE/P010660/1

2019: focus on deeper clinical and social phenotyping

1. Laboratory results: Wales Results Reporting Service (WRRS); 86M rows
2. Uncoded laboratory reports (text): also through WRRS with privacy protecting NLP MedGate architecture solution
3. Uncoded health board clinical correspondence: MedGate
4. National dispensing data with NHS Business Services
5. Local authority data flows through ADRC Wales
6. Images: radiology, pathology and retinal photographs
7. Building links with ONS and HMRC – ADRC Wales

The new UK Administrative Data World – 2018-2021+

Admin Data Research Partnership – (ADRP)



The old world (ADR v1)



Admin Data V1

Researcher driven

User support

Safe settings

Data access negotiation

Centralised administration

The new world (ADR v2)

Administrative Data Research Partnership (ADRP)



Admin Data V2

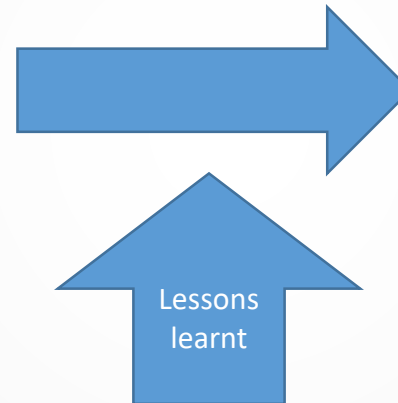
Impactful Policy projects

Government partners

Defined research programmes (SIPs)

UK strategic approach

Available data resources



Lessons learnt

Know what 'success' looks like and to whom

Have a defined data offering at the commencement of the programme.

Phase the service offering, particularly concerning data offer to researchers to manage expectations.

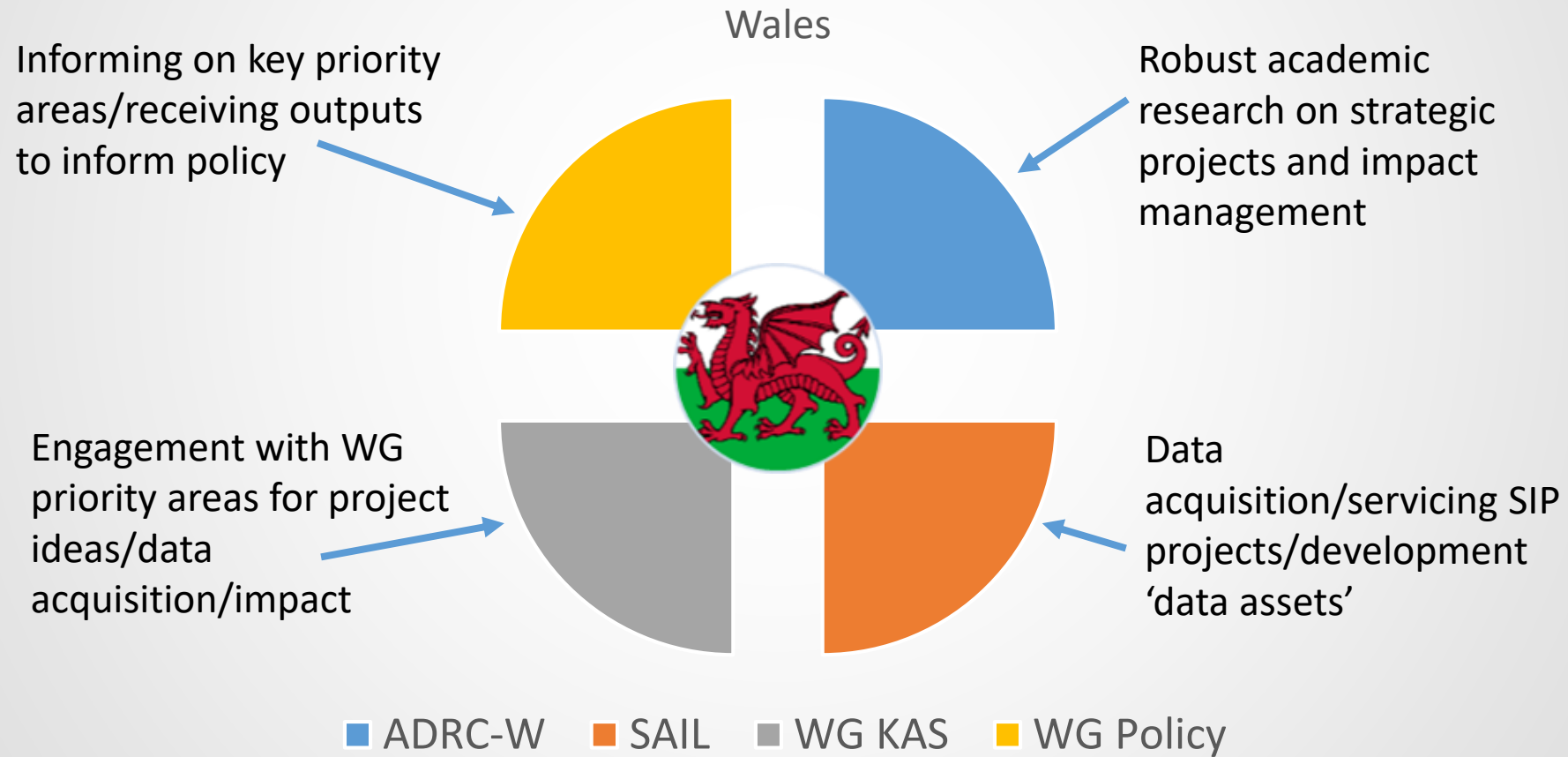
Define the operating model, processes and measurement Inc. success criteria at the planning stage of the project

Evaluate on a more proactive basis than mid-term reviews

Define benefits for data suppliers. Greater relationships with Government

Ensure geography is not a barrier to accessing services

Wales





Housing	Skills and employability	Mental Health	Wellbeing	Early years	Social Care	Emerging priorities
Academic lead (P Mackie)	Academic lead (C Taylor & R Davies)	Academic lead (A John)	Academic lead (R Lyons)	Academic lead (S Brophy)	WG lead	G Jones
Policy lead	Policy lead	Policy lead	Policy lead	Policy lead	Policy lead	
Research resource	Research resource	Research resource	Research resource	Research resource	Research resource	
KAS	KAS	KAS	KAS	KAS	KAS	
Administration support Statistical support Public engagement Impact management						

Injury research

Multi-sourced injury surveillance

Evaluation of service re-configuration

EMRTS Cymru

Natural experiments around interventions

Care and Repair Cymru

Fire and Rescue Service Wellbeing interventions

The population burden of disease

Original article
6
OPEN ACCESS

All Wales Injury Surveillance System revised:
development of a population-based system to
evaluate single-level and multilevel interventions

Ronan A Lyons,^{1,2} Samantha Turner,¹ Jane Lyons,¹ Angharad Walters,¹
Helen A Snooks,¹ Judith Greenacre,² Ciaran Humphreys,² Sarah J Jones²

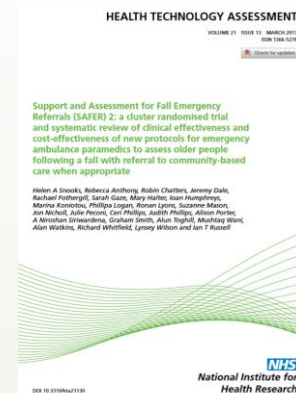
In Press: first published as 10.1136/

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

A Pooled Analysis of Vitamin D Dose
Requirements for Fracture Prevention

Heike A. Bischoff-Ferrari, M.D., Dr.P.H., Walter C. Willett, M.D., Dr.P.H.,
Endel J. Orav, Ph.D., Paul Lips, M.D., Pierre J. Meunier, M.D.,
Ronan A. Lyons, M.D., M.P.H., Leon Flicker, M.D., John Wark, M.D., Ph.D.,
Rebecca D. Jackson, M.D., Jane A. Cauley, Dr.P.H.,
Haakon E. Meyer, M.D., Ph.D., Michael Pfeifer, M.D., Kerrie M. Sanders, Ph.D.,
Hannes B. Stahelin, M.D., Robert Thiele, M.D., and Bess Dawson-Hughes, M.D.



JECH Online First, published on January 13, 2014 as 10.1136/jech-2013-203427

Research report

6
OPEN ACCESS

The association between hospitalisation
for childhood head injury and academic performance:
evidence from a population e-cohort study

Belinda J Gabbe,^{1,2} Caroline Brooks,¹ Joanne C Demmler,¹ Steven Macey,¹
Melanie A Hyatt,¹ Ronan A Lyons^{1,2,3}

J Epidemiol Community Health: first

PLOS MEDICINE

Browse | Publish | About

OPEN ACCESS | PEER-REVIEWED

RESEARCH ARTICLE

Long-term health status and trajectories of seriously injured
patients: A population-based longitudinal study

Belinda J. Gabbe, Pam M. Simpson, Peter A. Cameron, Jennie Ponsford, Ronan A. Lyons, Alex Collie, Mark Fitzgerald,
Rodney Judson, Warwick J. Teague, Sandra Braaf, Andrew Nunn, Shanthi Ameratunga, James E. Harrison

Published: July 5, 2017 • <https://doi.org/10.1371/journal.pmed.1002322>

VIBES Project

- Pooled longitudinal injury cohort data from US, UK, Netherlands, Australia, and New Zealand
- Aims to provide valid estimates of the burden of non-fatal injury using empirical data

PLOS ONE

Downloaded from <http://injury prevention.bmj.com/> on March 24, 2015 - Published by group.bmj.com

Disability weights in multinational injury studies

Belinda J Gabbe,^a Ronan A Lyons,^a Suzanne Polinder,^e Sarah Derrett,^f James E Harrison,^g Frederick P Rivara,^h Damien Jolley,ⁱ Shanthi Ameratunga,^j Suzanne Polinder,^k Juanita Haagsma,^k

Objective To create patient-based disability weights as an alternative to panel-based weights.

Methods Self-reported data based on the Injury-VIBES injury cohort study, which included participants from Ireland and the United States of America and for each type of diagnosis covered by the study. Weights were calculated separately for each type of diagnosis.

Findings There were 29 770 injury cases included in the study. Most participants were male and all cases were confined to emergency department presentations. Long-term disability was common in the study population.

Conclusion Injury is often a chronic condition and injury studies would substantially increase the burden of injury on peoples' lives.

OPEN ACCESS

Citation: Gabbe BJ, Simpson PM, Lyons RA, Ameratunga S, Harrison JE, Rivara FP, Jolley D, Derrett S, Polinder S, Haagsma J (2014) Disability weights in multinational injury studies. *PLoS ONE* 9(11): e111111. doi:10.1371/journal.pone.0111111

Editor: Hemachandra Iyer, University Health of America

Received: May 1, 2014

Accepted: October 1, 2014

Published: December 1, 2014

Copyright: © 2014 Gabbe et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

RESEARCH ARTICLE

Association of Injury and Outcome

How well do principal diagnosis classifications predict disability 12 months postinjury?

Validating and Improving Injury Burden Estimates Study: the Injury-VIBES study protocol

Belinda J Gabbe,¹ Ronan A Lyons,² James E Harrison,³ Frederick P Rivara,⁴ Shanthi Ameratunga,⁵ Damien Jolley,¹ Suzanne Polinder,⁶ Sarah Derrett⁷

ABSTRACT

Background Priority setting, identification of unmet and changing healthcare needs, service and policy planning, and the capacity to evaluate the impact of health interventions requires valid and reliable methods for quantifying disease and injury burden. The methodology developed for the Global Burden of Disease (GBD) studies has been adopted to estimate the burden of disease in national, regional and global projects. However, there has been little validation of the methods for estimating injury burden using empirical data.

Objective To provide valid estimates of the burden of non-fatal injury using empirical data.

measuring disease burden, the Disability Adjusted Life Year (DALY).^{6,7}

The DALY is a summary measure of population health which combines years of life lost (YLL) and years lived with disability (YLD).^{1,3,8} The YLL component uses incident mortality data. Calculation of YLD uses data related to incidence, prevalence and duration of disability attributable to particular conditions, combined with a disability weight for the specific disease or condition.⁹ Although the DALY metric is widely used, validation based on empirical data has been limited,² and we have previously raised limitations of the metric for estimating burden of injury in particu-

metric sums years of life lost (YLL) and the YLD component requires data about prevalence and duration of disability attributable to particular injuries, combined with a disability weight for the specific injury.² Disability weights represent injury diagnosis groups and do not reflect homogeneity in disability outcomes. A disability weight does not reflect all injuries in a group, overestimation or underestimation of injury will result. The diversity in severity of injuries sustained provides a challenge for grouping injuries into injury groups. Groupings used previously in injury studies have ranged from the 13

MONASH University

- VIBES was largely adult focused but 1/3 of the world's population are children and adolescents
- Childhood injury can adversely affect a child's prospects for health, education and social inclusion, as well as family and societal impacts
- Current burden of injury and disease methods fail to address differences between adults and children
 - Generic disability weights
 - Generic nature of injury classifications
 - Do not address the complexity of injury during key stages of growth and development
 - Limited long term outcomes data about child and adolescent injury to date

VIBES-Junior

- Funded by the National Health and Medical Research Council (2018-2020)
- **Are existing metrics for burden of injury based on adult populations valid for estimating burden of injury in children and adolescents?**
- Multi-national project using longitudinal injury studies and whole-of-population cohorts to:
 - Compare and contrast findings between study types
 - Provided a comprehensive evaluation of the long-term impact of injury in childhood and adolescence

Acknowledgements and participating organisations



Longitudinal studies

Pool and harmonise data

Diagnosis classifications

Disability weight calculations

Modelling trajectories

Additional linkages

Whole-of-population cohorts

Parallel analysis

Harmonisation of items

Rates of injury

Health care utilisation and other outcomes

Injured vs. uninjured

Welsh Electronic Cohort for Children (WECC)

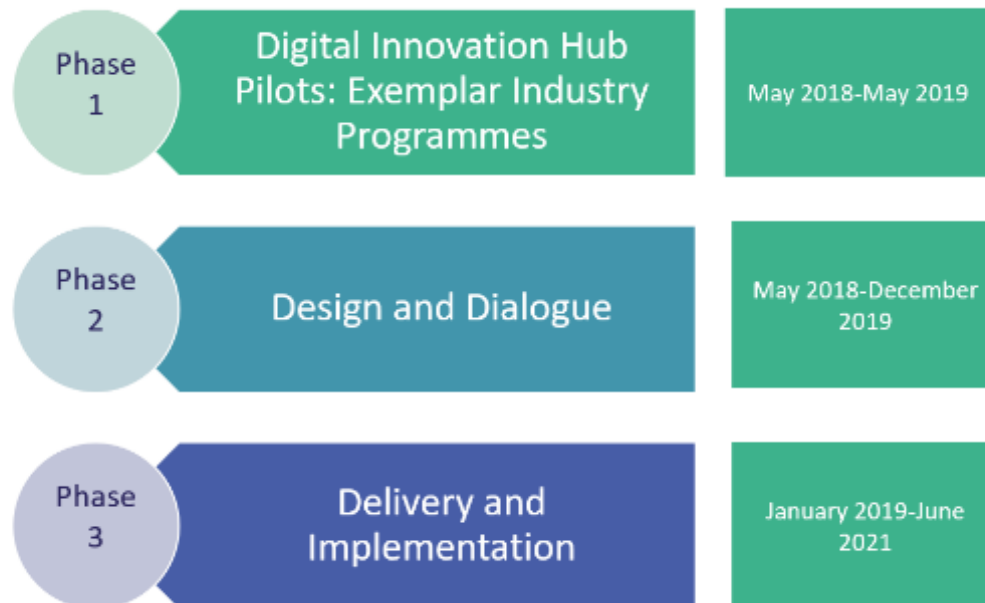
- Children born or living in Wales from 1 Jan 1998 to 31 Dec 2017
- Datasets
 - GP
 - Hospital inpatient and outpatient episodes
 - Emergency department presentations
 - Critical care minimum dataset
 - National Pupil Database
 - Pupil Level Annual School Census
 - Trauma Audit and Research Network
 - Burns datasets

South Australia Early Childhood Data Project

- Children born in South Australia from 1991 to 2016
- Datasets
 - Birth and pregnancy
 - Birth defects
 - Community maternal and child health services
 - Child development
 - Pre-school
 - Educational attainment
 - Hospitalisations and ED presentations

UK wide expert research data services

- Gerry Reilly joins as CTO
- Guiding principles for HDR UK national health data research infrastructure
- MRC £10M capital opportunity
- Blueprint of national strategy by September 2018
- Links to Local Health and Care Record Exemplars (LHCREs)



UK Secure e-Research Platform (UKSeRP)

Federated Infrastructure
Unified User Experience
Changing the status quo

Simon Thompson, CTO, Swansea University Medical School





Medical School
Ysgol Feddygaeth



Llywodraeth Cymru
Welsh Government



biobank^{uk}
Imaging study



Dementias
Platform^{UK}
Medical Research Council

MS*
REGISTER



UKSeRP Tenancies



Remote Access	Security	Data Management	Data Storage	Administrative Control	File Storage
Data Analysis Tools	BioInformatics	MedGate (NLP)	Imaging	Genomics	Technical Support

- Public-private consortium (MRC, universities, industry)
- Swansea provides remote analysis platform and data support services
- Network brings together 56 cohort studies from across the UK and internationally (45 UK, 5 South Korea, 3 USA, 1 France, 1 Ireland, 1 Hong Kong)
- Population or disease cohorts looking at cognition, ageing, or dementia specifically or as major themes
- Governance completely retained by cohorts
- Centralised rapid data access (analysis environment), central security systems (S3), supported by data curation, management, categorisation

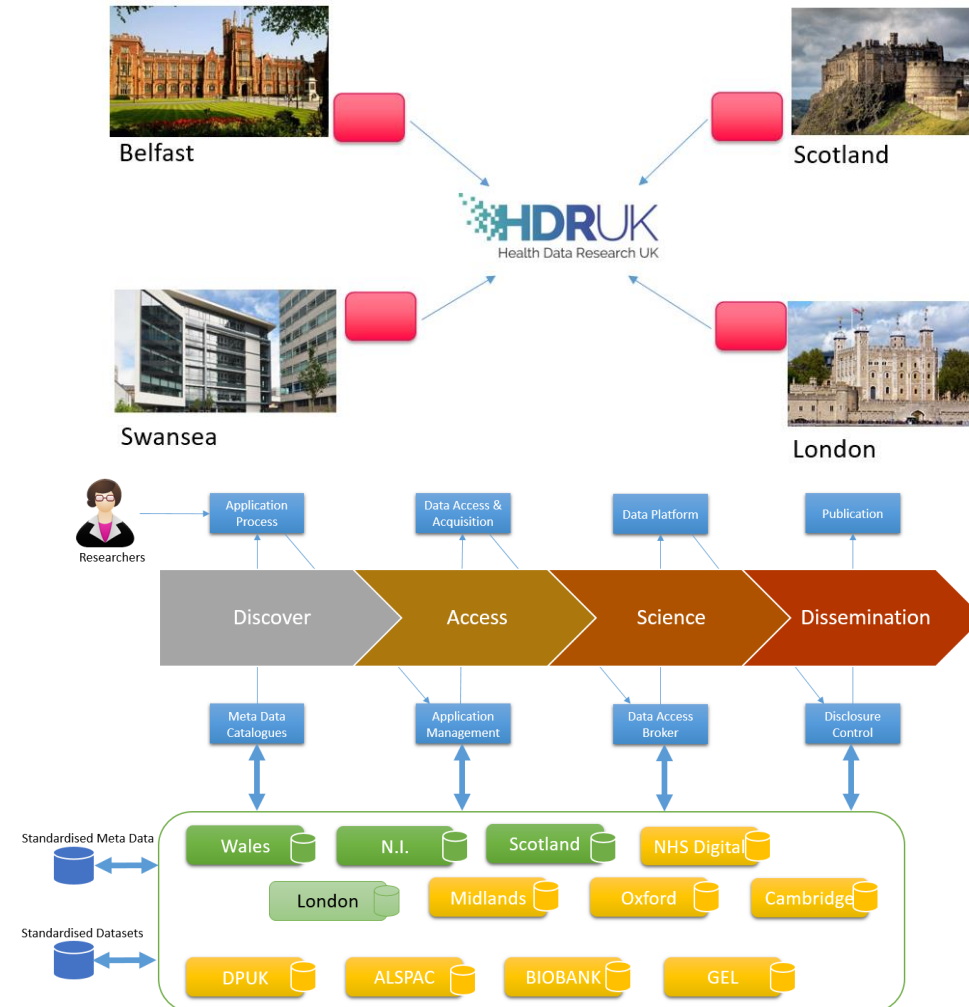
One UK Federated Research Network

Enhancement through collaboration

Challenge Areas

- Distributed analysis / query
- Data Harmonisation
- Data Discovery & Documentation
- Data Visualisation
- Analytics Platforms
- Multi Modality Analysis
- Enterprise/Cyber Security
- Compliance and Audit
- Data storage and archive
- Scalability / Affordability

POC Help Inform Design / Requirements



Liberating image data

- Multiple sources
 - Radiology
 - Pathology
 - Photos (retinal screening)
 - Linkable to EHR
 - NLP derived phenotypes
 - Privacy protection – image ID removal/modification

Exploratory investigation and visualisation of the relationships between multi-morbidity, polypharmacy, frailty and health service utilisation in Wales, UK

Jane Lyons, Damon Berridge, Ashley Akbari, Joe Hollinghurst, Christina Meggs, Ronan Lyons

- 12th September 2018
- IPDLN, Banff

- Related but poorly defined concepts
- Increasing prevalence
- A global health challenge



Data linkage and visualisation have much to offer in understanding trends in prevalence, inter-relationships between variables and impact on healthcare activity and outcomes.

Aim

Investigate changing prevalence's of multi-morbidity, polypharmacy and frailty

Exploration of the relationship between concepts and scoring methodologies recorded in primary care with secondary care

Methods

Utilisation of the SAIL Databank to create an e-cohort of anonymised Welsh residents, living in Wales at a mid-year point and registered with a SAIL providing General Practice for at least 360 days of 2014 and calculated multi-morbidity indexes and reviewed the healthcare utilisation per person

Investigations

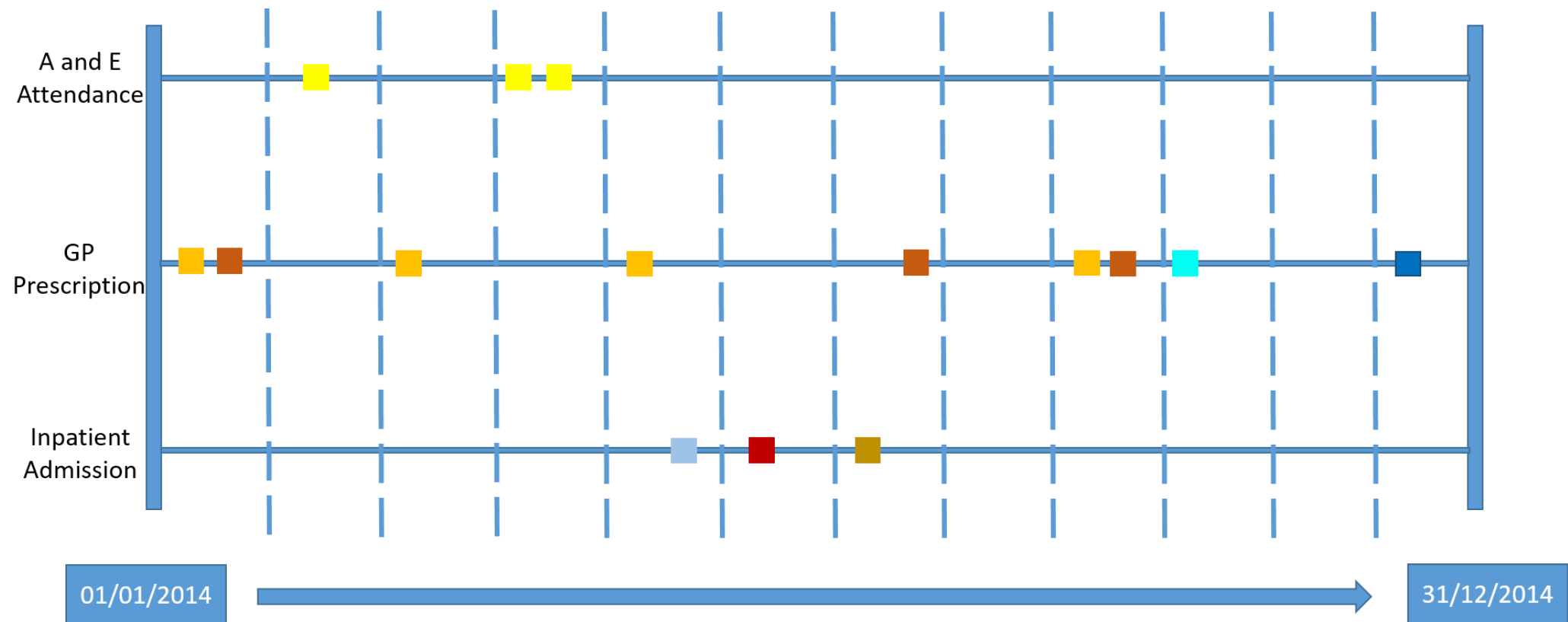
Calculation of two 'polypharmacy' scores in primary care

Calculation of multi-morbidity in primary and secondary care using 16 UK Quality and Outcomes Framework (QoF) and modified Charlson Index

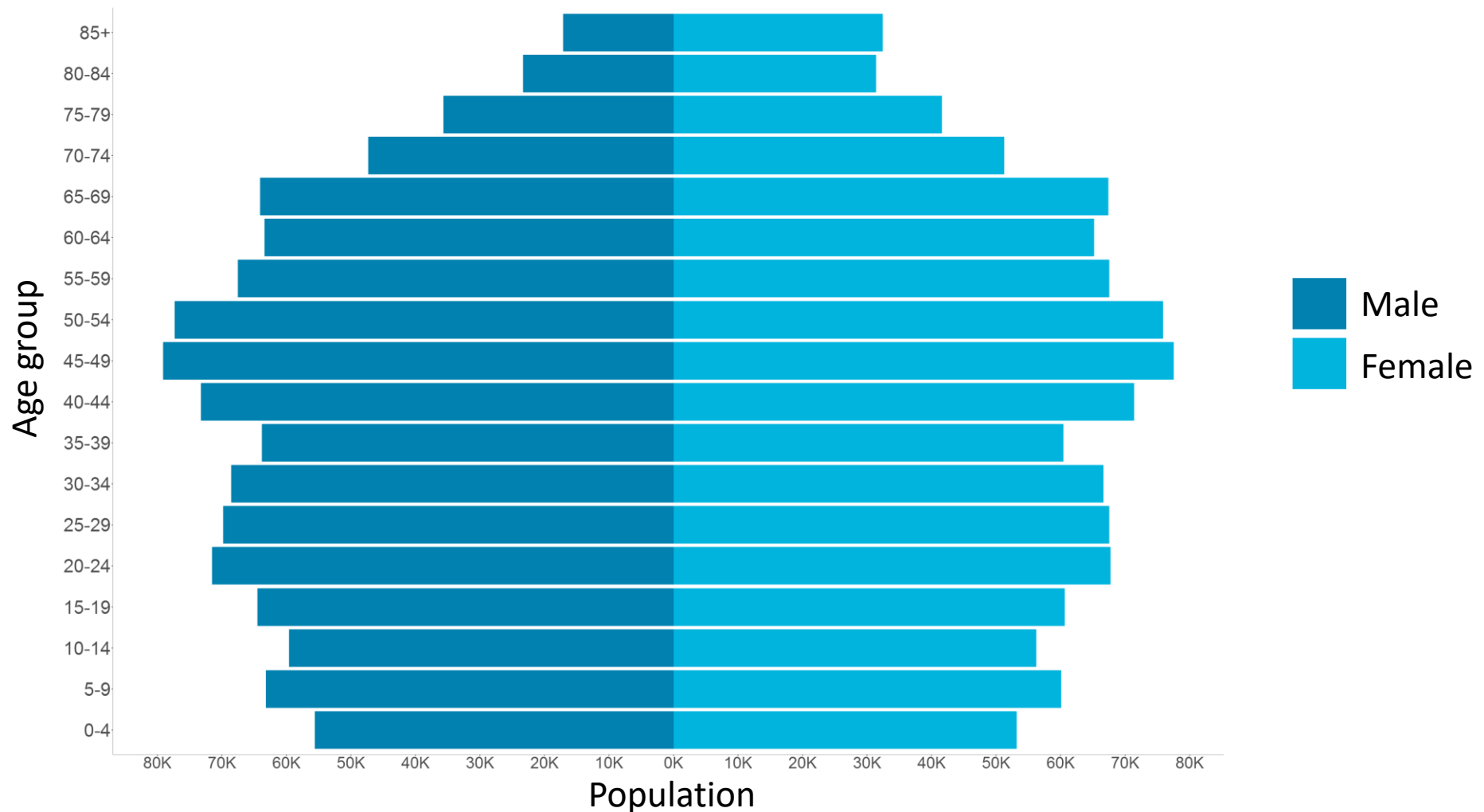
Implementation of Clegg's electronic frailty index in primary care (eFI)

Visualisation of the distributions of scores

Visualisation of relationships with number of outpatient appointments and GP polypharmacy

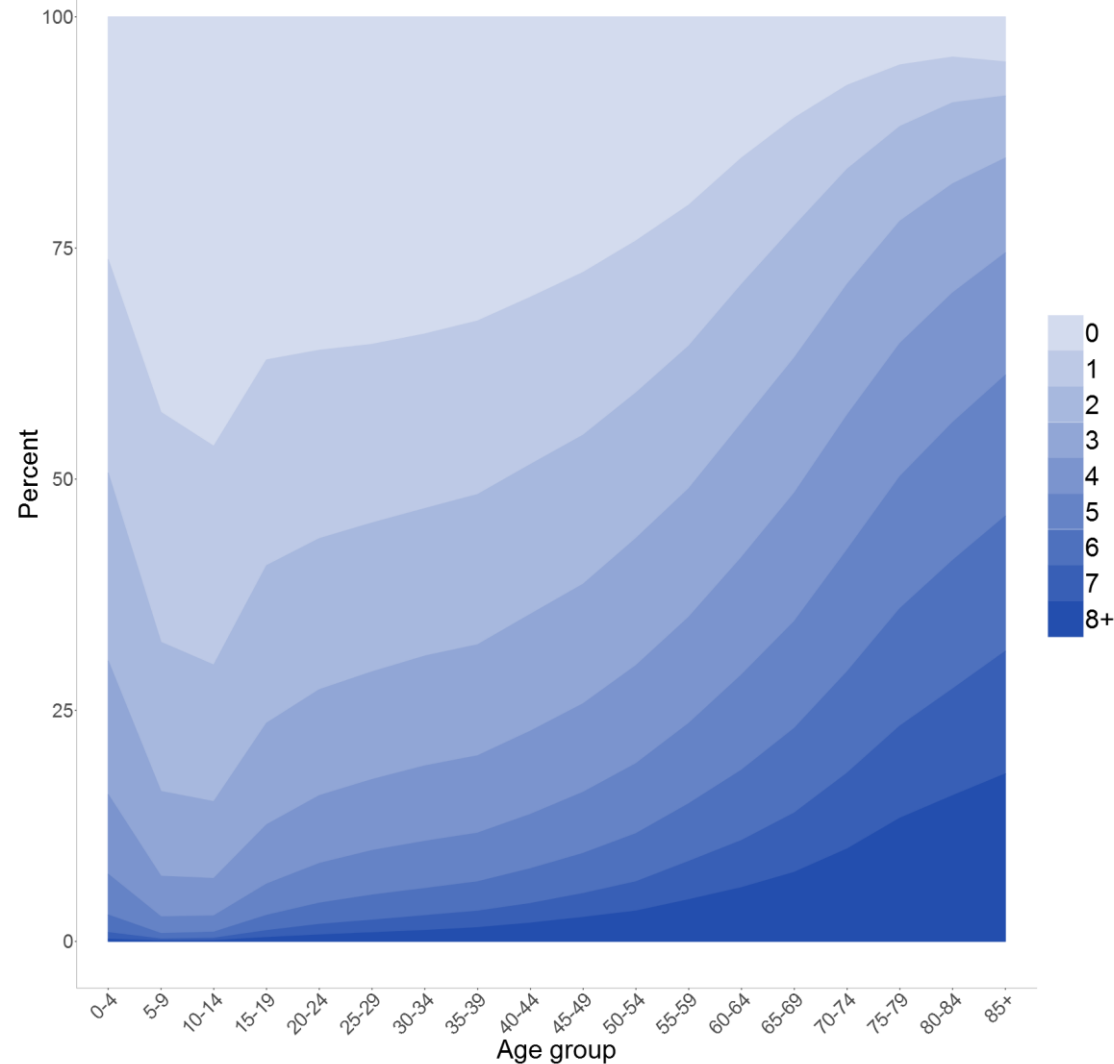


The e-cohort included **2.1 million** people. In 2014, **88.8%** had GP activity; **35%** an OPD episode; **18.7%** an ED attendance; and **14.2%** an admission

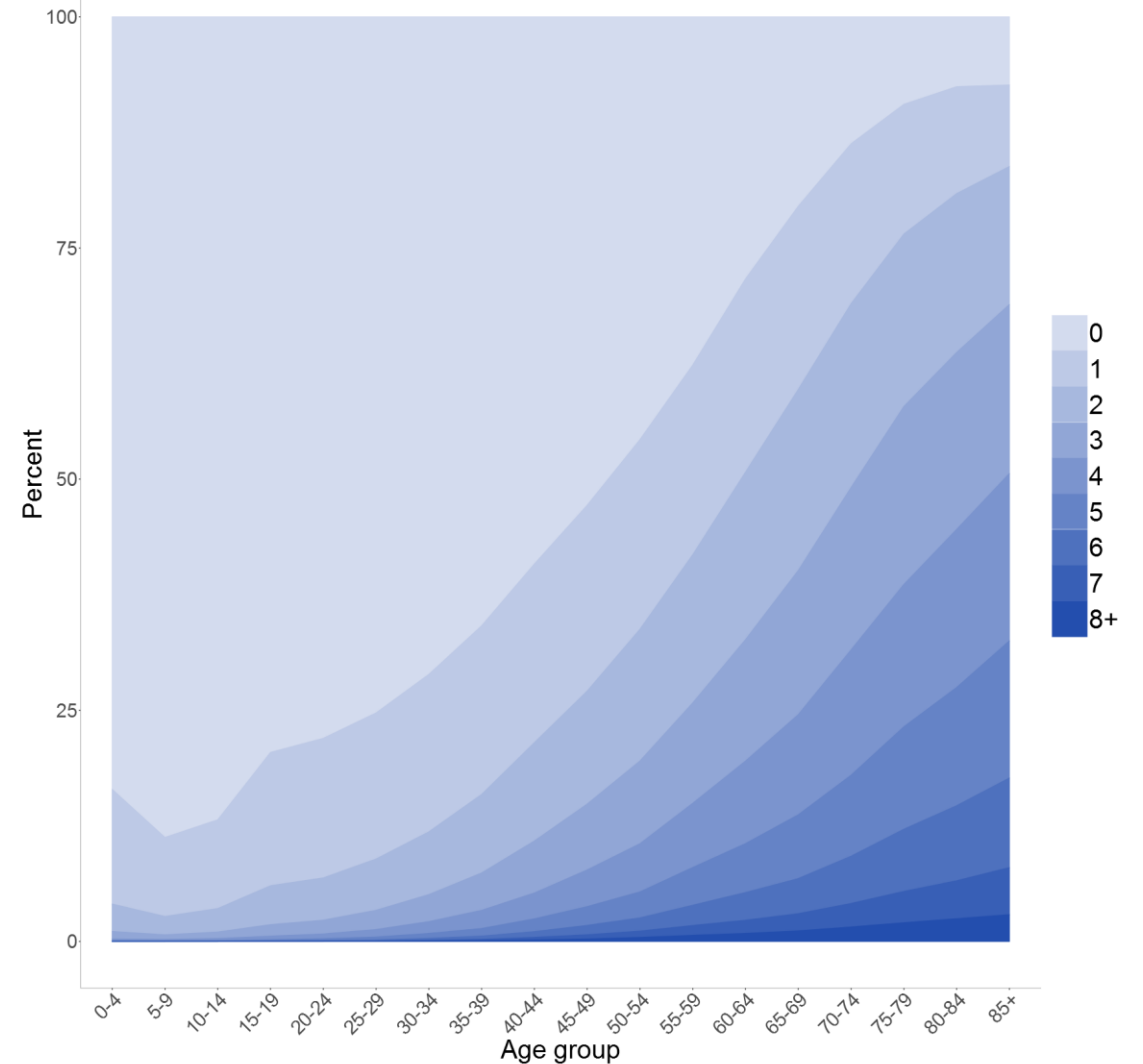


GP polypharmacy categories by prescribing threshold

Percentage of people in Wales by number of polypharmacy groups (any mention of 1 drug per Read code chapter) by age group, 2014



Percentage of people in Wales by number of polypharmacy groups (at least three distinct months of prescribing a drug per Read code chapter) by age group, 2014



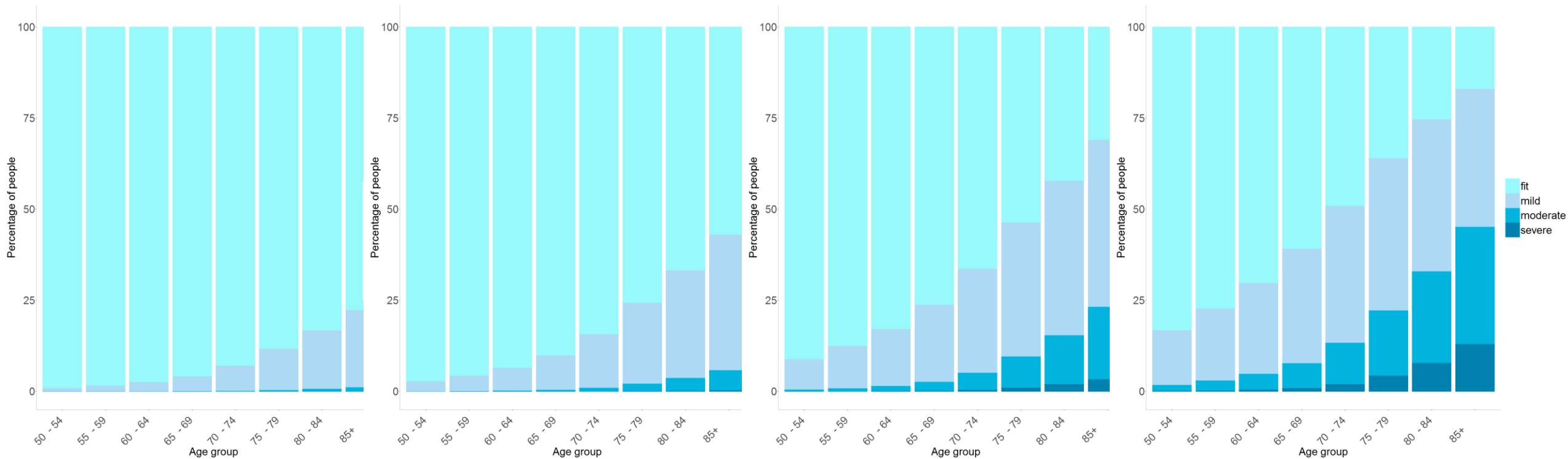
Proportion with eFI varies by duration of data look back (1-10 years)

The frailty index was calculated using primary care GP data and utilises 36 indicators to calculate the score.

Indicators include conditions such as AF, CVD, CKD, diabetes, dizziness, dyspnoea, falls, foot problems, fragility fracture, hearing impairment, heart failure, heart diseases, housebound, hyper/hypotension.

Each individual is initially scored 0 – 36 for any presence of an indicator in their GP data for a given time period. The total score per person is summed and divided by 36 to give a frailty index:
e.g. 9 indicators / 36 = 0.25 = Moderate

Frailty	Index
Fit	≤ 0.12
Mild	$> 0.12 \text{ \& } \leq 0.24$
Moderate	$> 0.24 \text{ \& } \leq 0.36$
Severe	> 0.36



Percentage of people in Wales by Frailty index and age group in 50 years and over, **2014**

Percentage of people in Wales by Frailty index and age group in 50 years and over, **2013-2014**

Percentage of people in Wales by Frailty index and age group 50 years and over, **2010-2014**

Percentage of people in Wales by Frailty index and age group in 50 years and over, **2005-2014**

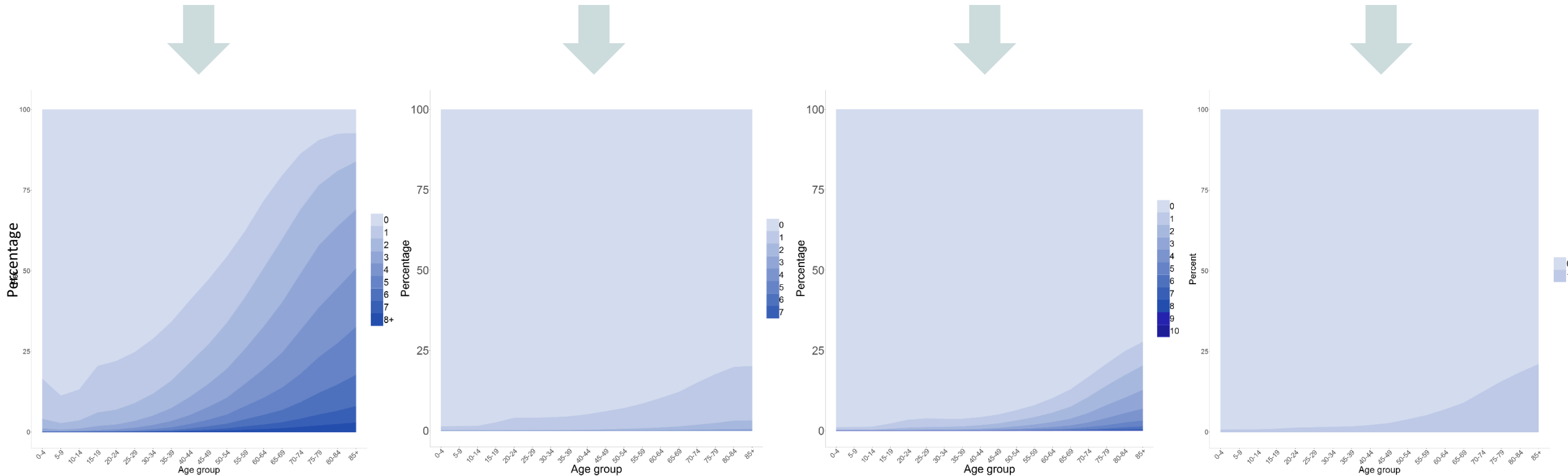
Results: Proportion of population flagged by different constructs in 2014

Primary care polypharmacy groups (with at least three distinct months of prescribing a drug per Read code chapter)

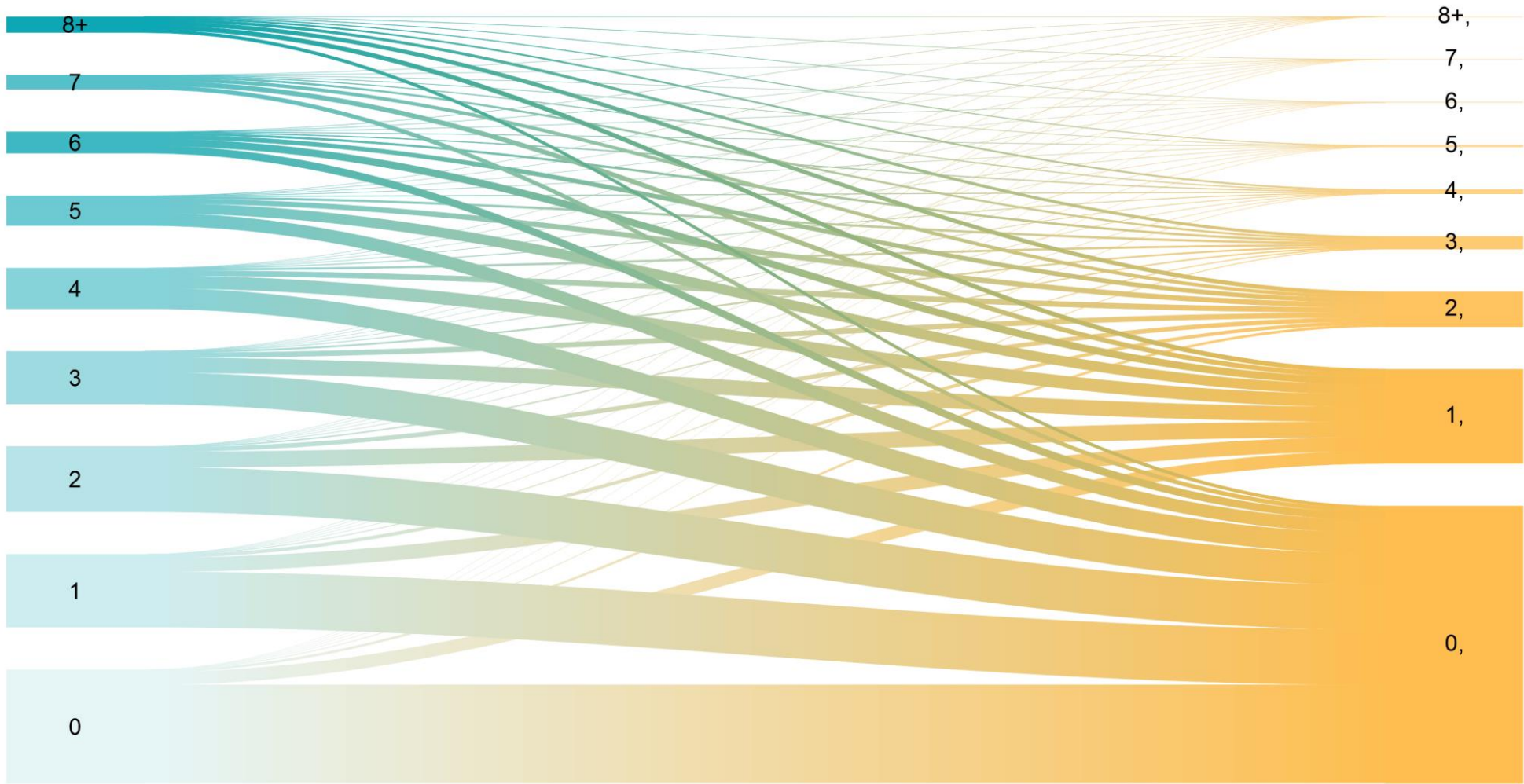
Primary care recorded multi-morbidity using 16 QoF indicators

Hospital admission multi-morbidity using 16 QoF indicators

Hospital admission multi-morbidity using weighted Charlson inpatient comorbidity score



Number of pharmacy chapter groups in GP data and the number of different outpatient clinics attended, 2014



Number of GP polypharmacy

Number of outpatient consultant specialties

- Proportion of population flagged by different constructs varies considerably
- Duration of look back strongly influences proportion flagged, especially in GP data
- Different individuals are flagged by different classifications: not 1 to 1 agreement
- Relationship between coded diagnoses, symptoms and prescribing not straight forward
- GP data better than in-patient for earlier identification of morbidity; much of polypharmacy not captured by standard morbidity indices
- Developments in agreed constructs of multimorbidity and longitudinal multi-sourced data linkage needed to produce greater insight into distribution and determinants of multimorbidity and allied constructs

Questions?

Contact information: R.A.Lyons@Swansea.ac.uk

HDRUK multi-morbidity implementation project

Investigating

- a. Patterns and trends
- b. Determinants and prevention
- c. Management

Federated analysis of geographically dispersed datasets utilizing team science approach

- Data from Wales, Scotland, NI, E London, Leicester....
- Wales: pop e-cohort with GP, hospital and laboratory data
- Supplement with new phenotypes derived from NLP of radiology, pathology and clinical correspondence
- Collaboration with Alan Turing Institute on Identification of 'novel' clusters of morbidity and their inter-relationships

Questions?

