Big data and cardiovascular disease research: opportunities in the Farr Institute

The Henderson Lecture

Welsh Cardiovascular Society
Cardiff 15 May 2014

Harry Hemingway FFPH, FRCP
Professor of Clinical Epidemiology, UCL
Director of Farr Institute, London

www.farrinstitute.org/
Wales and record linkages and registries for health

• On the **world stage**

• SAIL and now Farr Institute Ronan Lyons, David Ford and team [www.farrinstitute.org/](http://www.farrinstitute.org/)

• In cardiovascular Julian Halcox, Clive Weston
Vascular Risk Management in Wales

A report from the Vascular Project Group
March 2010

Together for Health – a Heart Disease Delivery Plan
A Delivery Plan up to 2016 for NHS Wales and its Partners

Delivery theme 5: Improving Information
Delivery theme 6: Targetting research
Cardiovascular diseases global #1

- Cause of death/premature death/disability adjusted life years
- So what has gone wrong?
  - wrong prevention
  - wrong treatments
  - wrong diagnoses / wrong names for diseases
  - wrong health systems (and too costly)
  - wrong relations to data, information and knowledge
  - wrong relations with patients
  - wrong science! (done by the wrong people!)
Might big data help right these wrongs?
Yes!

Mike Lauer, Director Division of Cardiovascular Sciences NIH-NHLBI

Transformation is in the Air...

Time for a Creative Transformation of Epidemiology in the United States

JAMA 2012
Yes!
Eric Topol ‘wireless and genomic medicine’
Big data

like teenage sex  ‘everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it’

Dan Ariely
What is big data?

**Volume**
- 40 Zettabytes (43 trillion gigabytes) of data will be created by 2020, an increase of 300 times from 2005.
- 6 billion people have a mobile phone.
- World population is 7 billion.

**Velocity**
- The New York Stock Exchange captures 1 TB of trade information during each trading session.
- Modern cars have close to 100 sensors that monitor items such as fuel level and tire pressure.

By 2015, 4.4 million IT jobs will be created globally to support big data with 1.5 million in the United States.

**Variety**
- It’s estimated that 2.5 quintillion bytes (2.5 zettabytes) of data are created each day.
- Most companies in the U.S. have at least 100 terabytes (100,000 gigabytes) of data stored.

**Veracity**
- As of 2011, the global size of data in healthcare was estimated to be 1.5 exabytes (1.5 billion gigabytes).
- As of 2014, it’s anticipated there will be 420 million wearable, wireless health monitors.
- 30 billion pieces of content are shared on Facebook every month.
- 400 million tweets are sent per day by about 300 million monthly active users.

**The FOUR V’s of Big Data**
- **Volume**: Different forms of data are generated from various sources.
- **Velocity**: Analysis of streaming data is crucial for real-time decision-making.
- **Variety**: The diversity of data sources and types.
- **Veracity**: The quality and accuracy of data.

**What is big data?**
- Big data encompasses information from multiple internal and external sources, such as social media, enterprise content, sensors, and mobile devices. It provides companies with the ability to analyze their products and services to better meet customer needs, optimize operations, and find new sources of revenue.

**Why is big data important?**
- It helps organizations make more informed decisions.
- It can lead to new business opportunities and innovations.
- It can improve customer experience and satisfaction.
- It can enhance predictive analytics and forecasting.

**Poor data quality costs the US economy $3.1 trillion a year.**

**27% of respondents**
- In one survey, 27% of respondents were unsure of how much of their data was inaccurate.
What is big data

Barabasi A. NEJM2007;357:404-7
Pace and scale of translation

Discoveries

Public health and clinical decisions ➔ health gain

Big data / Health records

Trials

Outcomes & quality research
Clinical data collected at the time of angiography in 1996/7

paper based, moth eaten, non-standardised, guessed (digit preference), lost, illegible, printed out from different ‘electronic systems’ etc..

new and untrumpeted linked to forerunner of Hospital Episode statistics for outcomes
Discovery
Genomics

500k participants, 47 baseline biomarkers and custom gene array data available in 2014, cardiac and brain imaging in 100k underway

Open access

Scalable approaches to disease phenotypes (startpoints or endpoints) based on linked electronic health record resources

- cardiac
- diabetes
- stroke
- cancer

Example of Farr Institute working across Wales, Scotland and England
Genetic variants identified from genome wide association studies:
   Phenotype largely not from EHR,
   Studied one at a time....

Do these associations replicate on phenome wide association studies using EHR?
EHR and genomics for discovery

Discovering new risk factor associations: CVD aggregates vs specific diseases
Are the risk factors the same?
To answer this question reliably we need

- **Scale:** e.g. >1 million adults followed for 5 years

- **Phenotypic resolution:**
  - Baseline risk factors
  - Follow up for disease outcomes

Cost to research funder of such data collection?
£0.00
The research costs are substantial

Information governance
Store, share, harmonise, analyse EHR data.....with scalable tools

And develop pool of clinical expertise
Multiple Record Linkages...needs expansion across NICOR registries

The CALIBER platform
Four nationwide EHR sources linked

Patient’s experience

Healthy, GP registration

Stable angina

Pneumonia hospitalization

Myocardial infarction hospitalization

See GP for follow-up

Death

Primary care

New patient check: blood pressure, smoking status, alcohol use etc.

Diagnosis of stable angina. Blood tests (e.g. cholesterol). Prescription of aspirin, nitrates etc.

Diagnosis of myocardial infarction

Blood tests, blood pressure. Prescriptions of beta blocker, statin, ACEi etc.

Sudden death

Hospitalization (HES)

Admit / discharge dates.
Primary diagnosis: Viral pneumonia, not elsewhere classified

Admit / discharge dates.
Primary diagnosis: Acute myocardial infarction
Procedure: Percutaneous coronary intervention

Disease Registry (MINAP)

ECG, cardiac markers.
Diagnosis: STEMI

Death Census (ONS)

Date of death. Cause:
1) Rupture of abdominal aortic aneurysm
2) Old myocardial infarction

What does linked record data look like?
...need tools
How to define phenotypes using multiple EHR data sources?
1001, 2000-01-01, af_gprd=1
1231, 2012-03-03, af_hes=3
1121, 2013-05-04,
af_procs_gprd=1
1511, 1993-01-11,
heart_valve_gprd=2
9913, 2012-05-21, af_hes=1
67222, 1994-08-11, af_hes=1
682444, 1993-01-01,
heart_valve_hes=2
af=1,
af_diag_source="primary care" af_diag_date=2001-12-01
AF algorithm
**Atrial Fibrillation**

<table>
<thead>
<tr>
<th>Name</th>
<th>af</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter</td>
<td>Circulatory disease/Atrial fibrillation</td>
</tr>
<tr>
<td>Definition</td>
<td>Diagnosis of atrial fibrillation.</td>
</tr>
<tr>
<td>Data Type</td>
<td>Categorical</td>
</tr>
<tr>
<td>Data sources</td>
<td>GPRD, HES</td>
</tr>
<tr>
<td>Dictionaries</td>
<td>Read, ICD10, BNF, Free text</td>
</tr>
<tr>
<td>Authors</td>
<td>K. Morley (UCL), Shah A. (UCL), Patel R. (UCL), Liam Smeeth (LSHTM), R. Schilling (St Bartholomews &amp; The Royal London Hospital), R. Hunter (St Bartholomews &amp; The Royal London Hospital)</td>
</tr>
<tr>
<td>Agreed</td>
<td>01/02/2013 (Revision 1)</td>
</tr>
<tr>
<td>Category</td>
<td>Definition</td>
</tr>
<tr>
<td>1</td>
<td>Historic AF diagnosis</td>
</tr>
<tr>
<td>2</td>
<td>AF diagnosis inferred</td>
</tr>
<tr>
<td>3</td>
<td>AF diagnosis confirmed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source variables</th>
<th>Description</th>
<th>Source</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Atrial fibrillation diagnosis</td>
<td>Primary care</td>
<td>af_gprd</td>
</tr>
<tr>
<td></td>
<td>Atrial fibrillation diagnosis</td>
<td>Secondary care</td>
<td>af_hes</td>
</tr>
<tr>
<td></td>
<td>Atrial fibrillation procedures</td>
<td>Primary care</td>
<td>af_proc_gprd</td>
</tr>
<tr>
<td></td>
<td>Atrial fibrillation procedures</td>
<td>Secondary care</td>
<td>af_proc_opcs</td>
</tr>
<tr>
<td></td>
<td>AF medication</td>
<td>Primary care</td>
<td>af_drugs_gprd</td>
</tr>
<tr>
<td></td>
<td>warfarin or digoxin prescription</td>
<td>Primary care</td>
<td>af_warfarin_digoxin</td>
</tr>
<tr>
<td></td>
<td>Deep vein thrombosis</td>
<td>Primary care</td>
<td>dvt_gprd</td>
</tr>
<tr>
<td></td>
<td>Deep vein thrombosis</td>
<td>Secondary care</td>
<td>dvt_hes</td>
</tr>
<tr>
<td></td>
<td>Pulmonary embolism</td>
<td>Primary care</td>
<td>pe_gprd</td>
</tr>
<tr>
<td></td>
<td>Pulmonary embolism</td>
<td>Secondary care</td>
<td>pe_hes</td>
</tr>
<tr>
<td></td>
<td>ECG Text/Notes text mining</td>
<td>Secondary care</td>
<td>Algorithm</td>
</tr>
</tbody>
</table>
CALIBER Data Portal

- Online data discovery tool [caliberresearch.org](http://caliberresearch.org)
- Access to all CALIBER phenotypes, algorithms and implementation details and scripts (SQL, R, Stata)
  - 45 users, 4 institutions, 538 phenotypes, >15,000 clinical diagnostic codes curated
- Standardization
  - Frontend is ICD10, backend becoming SNOMED-CT, LOINC
- A community rather than a static resource
  - Researchers contribute phenotypes and algorithms
  - Other researchers validate/enhance/correct them
Engaging clinicians in EHR phenotyping

• ....is proving a ‘route in’ for many academically minded clinicians at early, mid and late career stages
...back to our question....
12 CVDs: common risk
Inverse, weak and strong...what’s the ‘risk factor’?

≈10 million person years follow-up

Rapsomaniki, George et al CALIBER 2013

--- = Composite CVD
Inverse, null, weak and strong... what’s the ‘risk factor’?

<table>
<thead>
<tr>
<th>Initial presentation of cardiovascular disease</th>
<th>Number of events</th>
<th>Hazard ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable angina</td>
<td>12232</td>
<td>1.62 (1.49, 1.77)***</td>
</tr>
<tr>
<td>Unstable angina</td>
<td>5286</td>
<td>1.53 (1.32, 1.76)***</td>
</tr>
<tr>
<td>Coronary disease not further specified</td>
<td>10012</td>
<td>1.58 (1.45, 1.73)***</td>
</tr>
<tr>
<td>Non-fatal myocardial infarction</td>
<td>15191</td>
<td>1.54 (1.42, 1.67)***</td>
</tr>
<tr>
<td>Unheralded coronary death</td>
<td>5101</td>
<td>1.43 (1.23, 1.65)***</td>
</tr>
<tr>
<td>Heart failure</td>
<td>13072</td>
<td>1.56 (1.45, 1.69)***</td>
</tr>
<tr>
<td>Arrhythmia or sudden cardiac death</td>
<td>3218</td>
<td>0.95 (0.76, 1.19)</td>
</tr>
<tr>
<td>Transient ischemic attack</td>
<td>10990</td>
<td>1.45 (1.31, 1.60)***</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>5643</td>
<td>1.72 (1.52, 1.95)***</td>
</tr>
<tr>
<td>Stroke not further specified</td>
<td>10106</td>
<td>1.64 (1.48, 1.81)***</td>
</tr>
<tr>
<td>Subarachnoid hemorrhage</td>
<td>1260</td>
<td>0.48 (0.26, 0.89) *</td>
</tr>
<tr>
<td>Intracerebral hemorrhage</td>
<td>2265</td>
<td>1.28 (1.02, 1.62) *</td>
</tr>
<tr>
<td>Peripheral arterial disease</td>
<td>10074</td>
<td>3.0 (2.8, 3.2)***</td>
</tr>
<tr>
<td>Abdominal aortic aneurysm</td>
<td>3051</td>
<td>0.46 (0.35, 0.59)***</td>
</tr>
<tr>
<td>Other death</td>
<td>56168</td>
<td>1.10 (1.05, 1.17)***</td>
</tr>
</tbody>
</table>

Shah et al CALIBER under review
Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people

Eleni Rapsomaniki, Adam Timmis, Julie George, Mar Pujades-Rodriguez, Anoop D Shah, Spiros Denaxas, Ian R White, Mark J Caulfield, John E Deenfield, Liam Smeeth, Bryan Williams, Aroon Hingorani, Harry Hemingway

Summary

Background The associations of blood pressure with the different manifestations of incident cardiovascular disease in a contemporary population have not been compared. In this study, we aimed to analyse the associations of blood pressure with 12 different presentations of cardiovascular disease. [A: we have added a study aim here. Please amend if you wish]

Methods We used linked electronic health records from 1997 to 2010 in the CALIBER (CAdiovascular research using Linked Bespoke studies and Electronic health Records) programme to assemble a cohort of 1.25 million patients, 30 years of age or older and initially free from cardiovascular disease, a fifth of whom received blood pressure-lowering treatments. We studied the heterogeneity in the age-specific associations of clinically measured [A: Ok?] blood pressure with 12 acute and chronic cardiovascular diseases, and estimated the lifetime risks (up to 95 years of age) and cardiovascular disease-free life-years lost adjusted for other risk factors at index ages 30, 60, and 80 years. This study is registered at ClinicalTrials.gov, number NCT01164371.
Higher resolution epidemiology: blood pressure and 12 cardiovascular diseases

Cohort N = 2 million adults, >100,000 disease events

Myocardial infarction

<table>
<thead>
<tr>
<th>Systolic BP (mmHg)</th>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;90</td>
<td>0.47</td>
<td>[0.44; 0.49]</td>
</tr>
<tr>
<td>90 to 99</td>
<td>0.61</td>
<td>[0.58; 0.63]</td>
</tr>
<tr>
<td>100 to 119</td>
<td>1.40</td>
<td>[1.37; 1.44]</td>
</tr>
<tr>
<td>120 to 129</td>
<td>1.67</td>
<td>[1.62; 1.73]</td>
</tr>
<tr>
<td>130 to 139</td>
<td>1.92</td>
<td>[1.85; 1.99]</td>
</tr>
<tr>
<td>140 to 159</td>
<td>2.28</td>
<td>[2.19; 2.36]</td>
</tr>
<tr>
<td>&gt;180</td>
<td>2.72</td>
<td>[2.59; 2.85]</td>
</tr>
</tbody>
</table>

Hazard ratio for 20 mmHg Increase in Systolic Blood Pressure

Abdominal aortic aneurysm

<table>
<thead>
<tr>
<th>Systolic blood pressure</th>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;90</td>
<td>0.66</td>
<td>[0.64; 0.69]</td>
</tr>
<tr>
<td>90 to 119</td>
<td>1.06</td>
<td>[1.00; 1.12]</td>
</tr>
<tr>
<td>119 to 129</td>
<td>1.07</td>
<td>[0.99; 1.16]</td>
</tr>
<tr>
<td>129 to 139</td>
<td>1.04</td>
<td>[1.03; 1.17]</td>
</tr>
<tr>
<td>&gt;139</td>
<td>0.98</td>
<td>[0.96; 1.01]</td>
</tr>
<tr>
<td>&gt;180</td>
<td>0.92</td>
<td>[0.83; 1.02]</td>
</tr>
</tbody>
</table>

Diastolic blood pressure

<table>
<thead>
<tr>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.86</td>
<td>[0.78; 0.95]</td>
</tr>
<tr>
<td>0.92</td>
<td>[0.87; 0.97]</td>
</tr>
<tr>
<td>1.00</td>
<td>[1.00; 1.02]</td>
</tr>
<tr>
<td>1.16</td>
<td>[1.13; 1.23]</td>
</tr>
<tr>
<td>1.33</td>
<td>[1.30; 1.35]</td>
</tr>
<tr>
<td>1.40</td>
<td>[1.37; 1.43]</td>
</tr>
<tr>
<td>1.42</td>
<td>[1.39; 1.45]</td>
</tr>
<tr>
<td>1.66</td>
<td>[1.63; 1.70]</td>
</tr>
</tbody>
</table>

New knowledge

...a challenge for experimental medicine

Confirms what we know from combining multiple expensive studies

Add resolution

CALIBER Rapsomaniki et al 2014, Lancet in press
Blood pressure and 12 cardiovascular diseases

CALIBER Rapsomaniki et al 2014, Lancet in press
‘Higher resolution’ approaches: implications

- Disease mechanism
- Trial design
- Screening and risk prediction
Discovery Trials
Developing informatics platforms for stratified trials

Rapid feasibility

EHR-based eligibility counts

Protocol

EHR data sources

Recruiting

EHR randomisation

UCLP eConsent

Following up & safety

Real-time outcome dashboards

Embedded eCRF
Thrombus Aspiration during ST-Segment Elevation Myocardial Infarction

Ole Fröbert, M.D., Ph.D., Bo Lagerqvist, M.D., Ph.D., Góran K. Olavcrona, M.D., Ph.D., Elmir Omerovic, M.D., Ph.D., Thorarinn Gudnason, M.D., Ph.D., Michael Maeng, M.D., Ph.D., Mikael Aasa, M.D., Ph.D., Oskar Angerås, M.D., Fredrik Calais, M.D., Mikael Danielewicz, M.D., David Erlinge, M.D., Ph.D., Lars Hellsten, M.D., Ulf Jensen, M.D., Ph.D., Agneta C. Johansson, M.D., Amra Käregren, M.D., Johan Nilsson, M.D., Ph.D., Lotta Robertson, M.D., Lennart Sandhall, M.D., Ivar Sjögren, M.D., Ollie Östlund, Ph.D., Jan Harnek, M.D., Ph.D., and Stefan K. James, M.D., Ph.D.

METHODS

We conducted a multicenter, prospective, randomized, controlled, open-label clinical trial, with enrollment of patients from the national comprehensive Swedish Coronary Angiography and Angioplasty Registry (SCAAR) and end points evaluated through national registries. A total of 7244 patients with STEMI undergoing PCI were randomly assigned to manual thrombus aspiration followed by PCI or to PCI only. The primary end point was all-cause mortality at 30 days.

RESULTS

No patients were lost to follow-up. Death from any cause occurred in 2.8% of the
Temporal resolution

...cardiologists can study both

Onset

Prognosis

Healthy ➔ Onset of first cardiovascular disease ➔ Second cardiovascular disease, death
**Prognosis research: opportunities in EHR**

**PROGRESS 1: A framework for researching clinical outcomes** *BMJ.* 2013 Feb 5;346:e5595.


**PROGRESS 4: Stratified Medicine Research** *BMJ.* 2013 Feb 5;346:e5793.
Outcomes research: capturing clinically meaningful complexity
one startpoint to many types of endpoint
Acute myocardial infarction: a comparison of short-term survival in national outcome registries in Sweden and the UK

Sheng-Chia Chung, Rolf Gedeborg, Owen Nicholas, Stefan James, Anders Jeppsson, Charles Wolfe, Peter Heuschmann, Lars Wallentin, John Deanfield, Adam Timmis, Tomas Jernberg, Harry Hemingway

Figure 3: Kaplan-Meier curves for cumulative mortality at 30 days after admission with acute myocardial infarction in Sweden and the UK
‘Real world’ prognosis of stable CAD (n=102,023) and 5 yr risk of coronary death + non-fatal MI (n=8,856)

A ‘gold standard’ for estimating relevant risks?

Rapsomaniki et al CALIBER EHJ 2013
Prognostic models using linked EHR: Which clinically recorded factors add to discrimination?

<table>
<thead>
<tr>
<th>Characteristics used in risk-assessment (incrementally updated models)</th>
<th>Difference in the C-index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. SOCIODEMOGRAPHICS</strong>&lt;br&gt;age, sex, deprivation</td>
<td>Diff in C-index 0.000 [0.000; 0.000] 0.703</td>
</tr>
<tr>
<td><strong>2. CAD DIAGNOSIS &amp; SEVERITY</strong>&lt;br&gt;CAD subtype, PCI/CABG, hist MI, nitrate use</td>
<td>0.042 [0.037; 0.047] 0.744</td>
</tr>
<tr>
<td><strong>3. PRIMARY CVD RISK FACTORS</strong>&lt;br&gt;smoking, hypertension, diabetes, lipids</td>
<td>0.009 [0.007; 0.011] 0.753</td>
</tr>
<tr>
<td><strong>4. CVD CO-MORBIDITIES</strong>&lt;br&gt;heart failure, PAD, atrial fibrillation, stroke</td>
<td>0.014 [0.013; 0.016] 0.758</td>
</tr>
<tr>
<td><strong>5. NON.CVD COMORBIDITIES</strong>&lt;br&gt;chronic renal disease, COPD, cancer, liver disease</td>
<td>0.001 [0.001; 0.002] 0.759</td>
</tr>
<tr>
<td><strong>6. PSYCHOSOCIAL</strong>&lt;br&gt;depression, anxiety</td>
<td>0.000 [0.000; 0.000] 0.769</td>
</tr>
<tr>
<td><strong>7. BIOMARKERS</strong>&lt;br&gt;heart rate, creatinine, white cell count, haemoglobin</td>
<td>0.009 [0.006; 0.012] 0.778</td>
</tr>
</tbody>
</table>

Origin of data is EHR therefore implementation of risk prediction models in decision support tools (with evaluation) is feasible.

Rapsomaniki et al CALIBER EHJ 2013
Discovery
Trials
Outcomes research/real world evidence
Public Health
Outcomes assessment: importance of linking multiple record sources

Death registry

Disease registry

Hospital admissions

Primary care

All four data sources

Crude annual incidence of myocardial infarction per 100 000

Herrett, Shah et al CALIBER BMJ 2013
How does CVD first present?

In the real world, today

- MI/Fatal CHD: 18%
- Unstable angina: 5%
- Ventricular arrhythmia/sudden cardiac death: 3%
- Abdominal aortic aneurysm: 2%
- Intracerebral haemorrhage: 2%
- Subarachnoid haemorrhage: 1%
- Ischaemic stroke: 13%
- Stable angina: 12%
- Heart failure: 12%
- Peripheral arterial disease: 11%
- Transient ischaemic attack: 11%
- CHD: 10%

N=1.93 million patients
>110K CVD events
5 year median follow-up

CALIBER 2014, under review
Strengthening health informatics research

- MRC coordinated 10-partner £19m call for e-health informatics research centres across the UK
- Cutting edge research using data linkage capacity building
- Additional £20m capital to create Farr Institute
- UK Health Informatics Research Network
  Coordinate training, share good practice and develop methodologies

Engage with the public, collaborate with industry and the NHS

- Farr London
- Farr Scotland
- Farr at Swansea, Wales
- Farr N8, Manchester
Rapid evolution of initiatives: emphasis on infrastructure

July 2013: Farr Institute of Health Informatics Research
July 2013: Genomics England
October 2013: Administrative Data Research Centre England
November 2013: National Institute for Health Research
February 2014: Medical Research Council

November 2013: Health Informatics Collaboration (sharing hospital data across 5 Biomedical Research Centres)

.....needs cardiovascular co-ordination

UCLPartners
Academic Health Science Partnership
Clinical data: 1996/7

paper based, moth eaten, non-standardised, guessed (digit preference), lost, illegible, printed out from different ‘electronic systems’ etc..

.....and in 2014 ?
Cardiology data – submerged in diverse “information” systems:
Drought

• **Data**
  – Need much wider national record linkages – CPRD-NICOR-HES
  – Need to liberate ‘submerged’ deeper hospital phenotypes
  – Need to converge EHR, omics and imaging

• **Tools**
  – Health informatics ‘20 years behind bioinformatics’

• **People**
  – (re) building **public trust** in the wake of care.data
  – Not nearly enough **clinicians** with the training and opportunity to drive improvements in care (and research) through data, information and knowledge (400 US clinicians in new board certified specialty JAMA 2014)
Farr and CALIBER cardiovascular collaborators

Mike Barnes, Mark Caulfield, Sheng-Chia Chung, John Deanfield, Spiros Denaxas, Chris Gale, Aroon Hingorani, Mar Pujades, Eleni Rapsomaniki, Anoop Shah, Liam Smeeth, Adam Timmis among others

among others
Wales, SAIL, CIPHER and the Welsh Farr node

- Wales is on the world stage in record linkages
- The funding environment is currently permissive (Wales representation at BHF/MRC/WT cardiovascular informatics workshop)
- So whatever stage of the translational cycle and informatics cycle you are interested in....

Get involved!

www.farrinstitute.org/