


CHIEF SCIENTIST OFFICE

Farr The Farr Institute of Health Informatics Research
 
NHS SCOTLAND

Options and Opportunities for Health Data Science

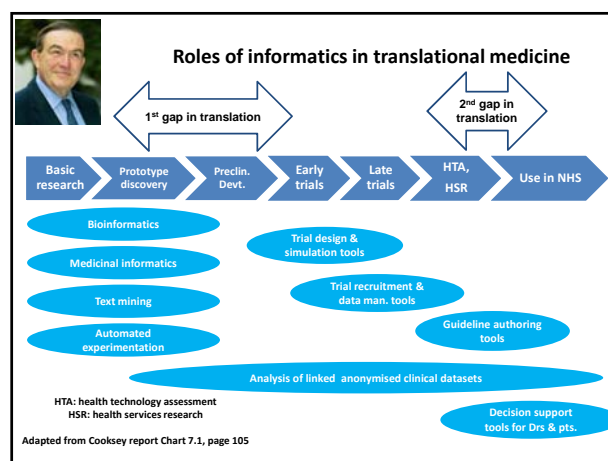
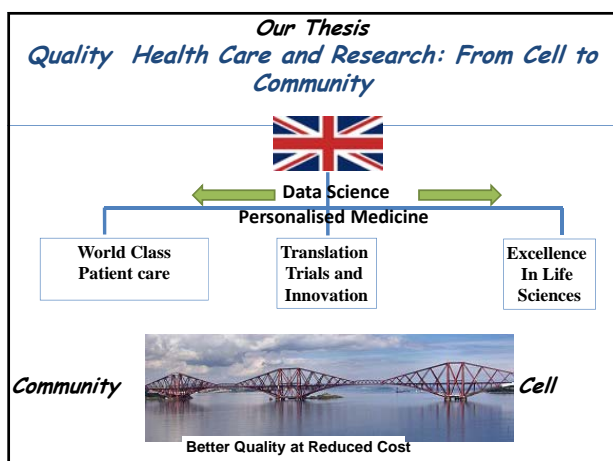

PERSONALISED MEDICINE CONFERENCE 2016
 1-2 June, Brussels

1st June 2016
 Andrew Morris
 Vice Principal Data Science, University of Edinburgh
 Chief Scientist Health Scottish Government
 Director Farr Scotland


PERSONALISED MEDICINE CONFERENCE 2016
 1-2 June, Brussels

The Next 15 minutes

- Gearing an entire country for quality health care and precision medicine
- Data Science as the catalyst for change
- Case studies of phenotypic definition
- With big data goes big responsibilities



The Future? "4P" Medicine

- **Predictive** *Customise diagnosis and treatment*
- **Pre-emptive** *Better than curative*
- **Personalised** *Determine risk profiles, predict outcomes*
- **Participatory** *Involve patients*

Made Possible by:

- Genomics
- Phenotyping
- Informatics
- Analytics
- New social contract



Informatics to support patient care



15.6 Billion Euros
Population 5M
Single health care provider
14 Territorial Boards
38 Hospitals, 1020 General Practices
High rates of morbidity of common complex disease
Collaboration – Aberdeen, Edinburgh, Dundee, Glasgow, St Andrews
Unique patient identifier



The Scottish Health Service on a Slide

	2012/13	5 year Change
Key Trends		
Population: 5.3 million		
% aged 75+ : 7.9%		
GDP Per Head in 2011: \$42,124		
Estimated GP Patient Contacts	16,539,000	3.3%
Estimated Practice Nurse Patient Contacts	7,627,000	10.5%
New A&E Attendances	1,561,529	6.8%
Total Outpatient Attendances	4,699,868	4.7%
Total Inpatient/Day Case Discharges	1,582,305	6.8%
Inputs		
• Acute Beds in 2011/12: 16,500 (NHS)		
• Doctors in 2012: 12,000 (NHS WTE)		
• Nurses / Midwives in 2012: 56,600 (NHS WTE)		
Day Case Discharges	448,782	10.6%
Routine Inpatient Discharges	441,024	9.5%
Non-Routine (emergency) Inpatient Discharges	540,890	6.4%

Urgent need to migrate from measurement of activity to REAL TIME MEASUREMENT of processes and outcomes meaningful for patients

Community Health Number

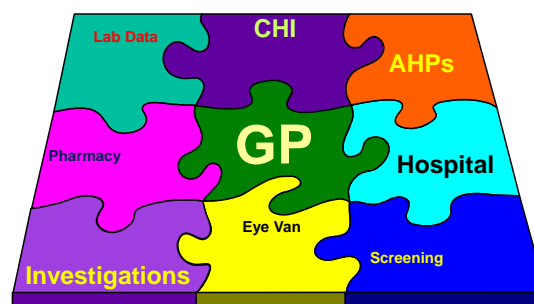
07 10 64 02 5 0

Date of Birth

Sex Check

Linking Data

the key to seamless care



National level data resources for 5M citizens

Neonatal Record		GP consultations		Substance misuse	
Mental Health		Community care			
BIRTH	Dental	Out patients	Hospital Admissions	DEATH	
Maternity	Prescribing	A&E	Screening	Suicide	Cancer registrations
Child health surveillance	Immunisation	Imaging	Laboratory		

Education		Marriage		Care homes	
Looked after children		Community care			
BIRTH	HMRC	DWP	Census (Scotland & UK)	DEATH	

Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- Driving clinical research
- Precision Medicine

Prevention of Disease Non-experimental evaluation (policy) Effect of smoking legislation

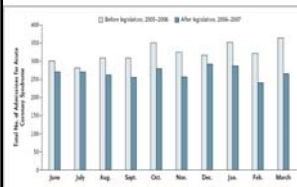


Figure 1. Admissions for Acute Coronary Syndrome According to Month before and after Smoke-free Legislation.

Admissions fell by 17% - 67% of reduction was in non-smokers
Fall in England 4% (no legislation); long term trend 3%

Acute Coronary syndrome

Pell et al, N Eng J Med (2008) 359: 482-491

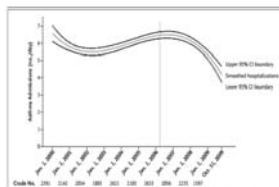


Figure 2. Daily Hospital Admissions for Asthma among Children between January 2008 and October 2009.

Before ban 5.2% increase per annum
After ban 18.2% decrease per annum

Childhood asthma

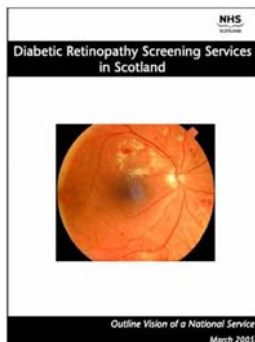
Pell et al New Engl J M 2010; 363, pp. 1139-1145

Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- Driving clinical research
- Precision Medicine

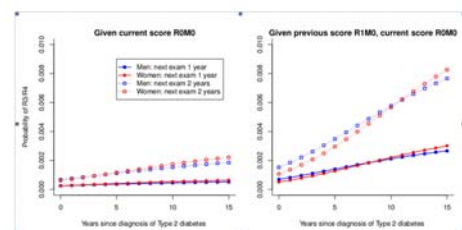
Diabetic Retinopathy Screening



Launched 2005
Annual screening interval
86% of eligible population screened (n=212,897 in 2014)
194 type 1 diabetes registered blind

Scottish Diabetes Survey 2014

Can we stratify to increase screening interval?



- 11,275 cases of referable retinopathy
- If people with T2DM and two examinations showing no visible retinopathy were offered two-yearly screening
- 44% fewer people need screened

Looker et al Diabetologia 56: 1716-25; 2013

Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- Driving clinical research
- Precision Medicine

THE LANCET Respiratory Medicine



Cardiovascular and neuropsychiatric risks of varenicline: a retrospective cohort study

Daniel Kotz, Wolfgang Viechtbauer, Colin Simpson, Omer C P van Schayck, Robert West, Aziz Sheikh



Investigating the safety of varenicline

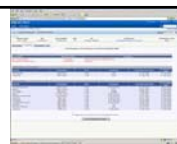
- Varenicline is by far the most effective smoking cessation treatment available
- Its use is however limited by an FDA warning of possible cardiovascular and neuropsychiatric side-effects
- Largest ever cohort study (~150k) using state-of-the-art methods linked to hospitalisation and mortality data, showed these concerns are unfounded
- Covered by Fox Business (amongst others), and Pfizer's shares have risen on the back of this....Likely to trigger FDA review...

Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- Driving clinical research
- Precision Medicine

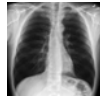
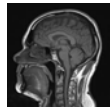
Emergency Care Summary



- Patient Safety is key driver
- Available nationwide
- Clinically Led, Patient Focused
- Twice daily updates from GP systems
- Medications and Adverse Reactions
 - Explicit Consent to view ECS
 - Full audit trail available at any time
 - Fully established 2006
 - Evaluation 2010
 - 3.5 Million Accesses per Annum
 - Changes management in 20%

A nationwide approach to digital radiology

- Common user interface across all hospitals
- 850TB of data
- 24,000 registered users in 38 hospitals
- Web-based image viewing in over 2000 wards
- 10% reduction in re-examination
- 100% reduction in film and chemical cost
- Significant reduced time to treatment
- Image database of 17 million studies for care and research



A National Diabetes System for Scotland



Total Scottish Population 5.2M

People with diabetes : 278,134 (5.1%)

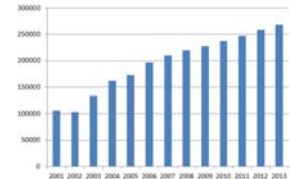
People with Type 1 DM : ~27,000 (0.5%)

Single clinical information system SCI-Diabetes

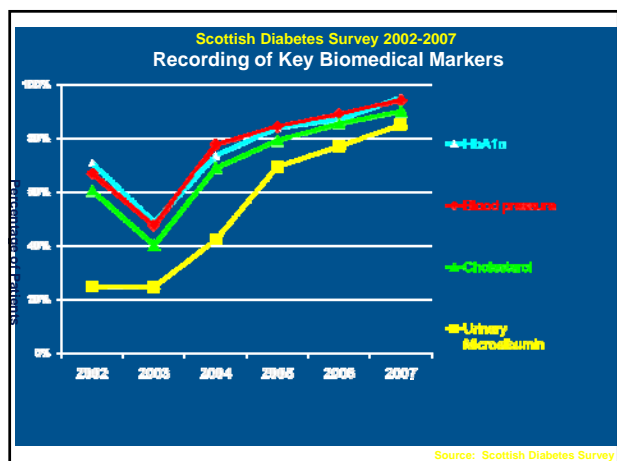
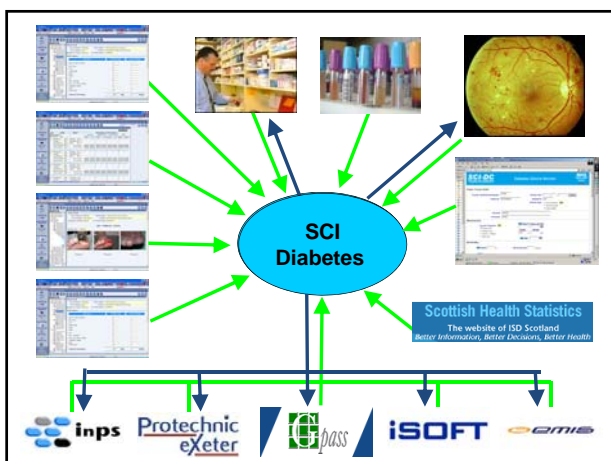
SCI-DIABETES used in all 38 hospitals

Nightly secure sharing of data from all 1043 primary care practices across Scotland

Linked DNA/longitudinal phenotype



Scottish Diabetes Survey 2014



Evidence of improved clinical outcomes

Original Article: Complications
Decreasing amputation rates in patients with diabetes—a population-based study

C. J. Schiffl, N. Yu*, A. S. Jain and G. P. Lane

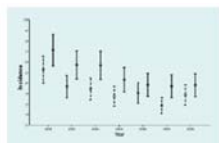
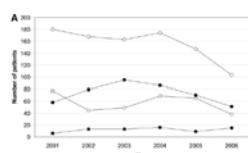


FIGURE 1 Incidence of adjusted total amputation with 95% confidence intervals (solid line) and incidence of adjusted major amputation with 95% confidence intervals (dashed line) per 1000 patients with diabetes, incidence adjusted for age and sex.

Diabetic Medicine 2009

Diabetic Retinopathy: More Patients, Less Laser

A longitudinal population-based study in Tayside, Scotland



Diabetes Care 2008

The Innovative Healthcare Delivery Programme

The Scottish Cancer Intelligence Framework Cancer Pathways

- Support patient care through an integrated cancer pathway
- Monitor clinical outcomes and effectiveness of treatment
- Electronic information follows the patient across organisational boundaries
- Maximise the use of data in existing clinical systems
- Adhere to relevant clinical data standards such as the Cancer Outcome and Services Dataset (COSD)



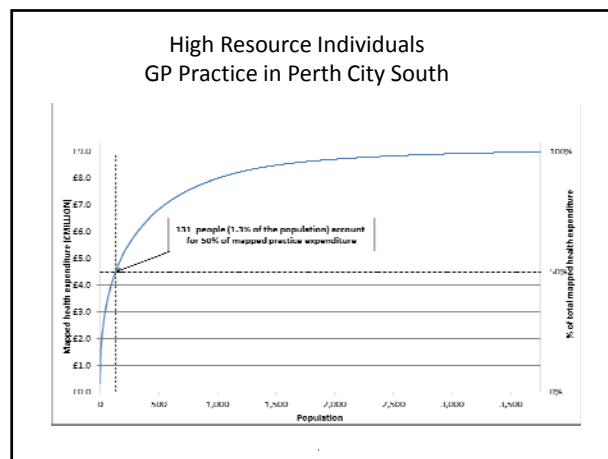
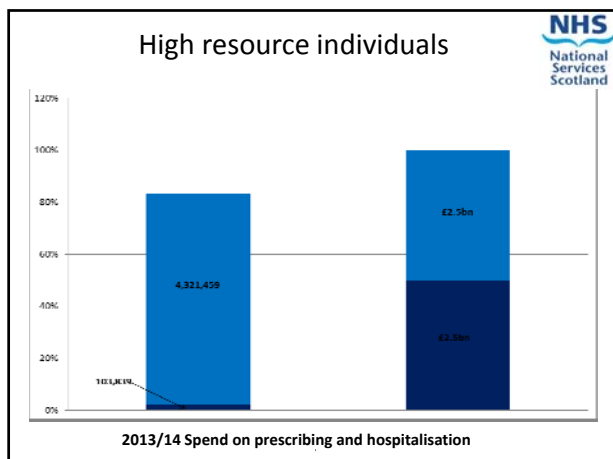
Pathway for Mr Smith



Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- Driving clinical research
- Precision Medicine



Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- **Driving clinical research**
- **Precision Medicine**



NHS
National Institute for
Health Research

Data Linkage Driving Efficiency in Clinical Trials



SHARE
Recruiting for Research, Improving our Health

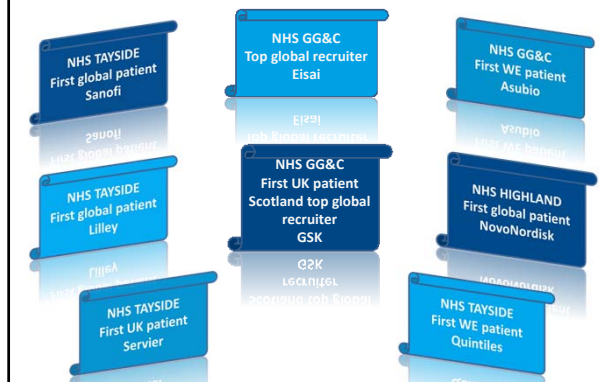
113,320 patients registered



Join dementia research

18,709 Volunteers
5,178 Volunteers Enrolled in Dementia Studies

Good News – 'UK First'

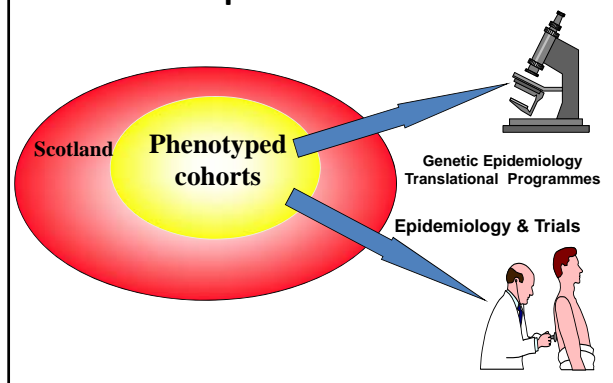


Case Studies



- Prevention of disease
- Earlier diagnosis
- Safer and more effective treatments
- More effective integrated pathways
- A greater understanding of how to use NHS funding most effectively
- Driving clinical research
- **Precision Medicine**

The Population Model



The next shift in technology

The first human genome sequence took 300 scientists, 15 years, and \$3bn



Now it takes 3 scientists, 1 day and \$1000

1%


0.02%

0.00003%

MRC Medical Research Council

Maximising the value of UK population cohorts

MRC Strategic Review of the Largest UK Population Cohort Studies

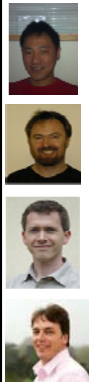


Adding Value to UK Cohort Studies

- 34 cohorts
- 2.5m people have taken part and currently around 2.2m people – 3.5% of the population – are cohort members.
- We owe them a debt of gratitude
- 500,000 people are part of UK Biobank and soon the entire cohort will be genotyped.
- Participants from the UK cohort studies have given consent for their personal data to be linked to NHS records and other data sources such as education and the census

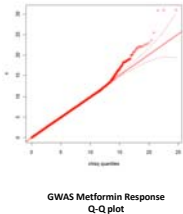
For Future of Research

Population Pharmacogenetics



- METFORMIN
- In use for over 50 years
- We still don't understand how it works
 - 25% of patients get GI intolerance;
 - 5% cannot continue it
- Can we use genetics to help us?
- Ability to link genetics with drug exposure and therapeutic response

wellcometrust



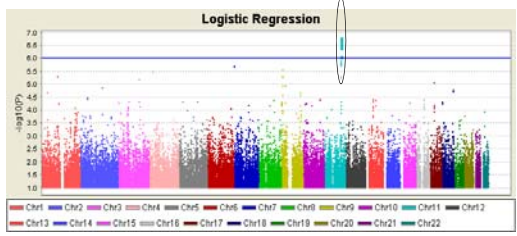
GWAS Metformin Response Q-Q plot

nature genetics

LETTERS

Common variants near *ATM* are associated with glycemic response to metformin in type 2 diabetes

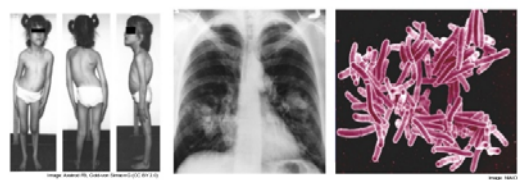
The GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group^{1,2} & The Wellcome Trust Case Control Consortium 2²



The gene links cancer pathways, metformin pathways and type 2 diabetes

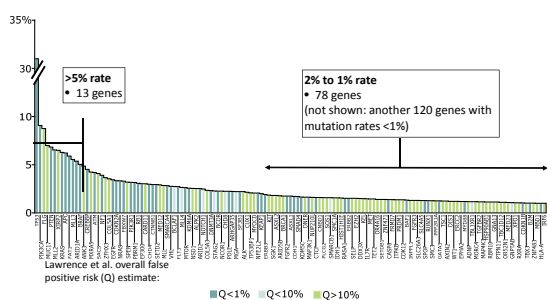
Genomics
england

100,000 genomes in 5 years



The need for an international perspective
Molecularly, every cancer is a rare cancer

Pan-cancer non-silent mutation frequency (%)



The need for an international perspective
Molecularly, every cancer is a rare cancer



Rank	Site	Incidence (per M)	Mortality	Millions of population for 250 incident cancer patients a year with biomarker at		
				20% frequency	5% frequency	1% frequency
1	Breast	653	23%	1.9	7.7	38
2	Prostate	548	26%	2.3	9.1	45
3	Lung	530	83%	2.4	9.4	47
4	Colon	340	37%	3.7	14.7	73
5	Melanoma	167	17%	7.5	29.9	149
10	Pancreas	111	94%	11.3	45.1	225
14	Ovary	87	62%	14.4	57.5	287
20	Liver	56	89%	22.4	89.5	447
25	Cervix	36	33%	34.5	138.2	690
30	Larynx	29	32%	43.3	173.4	866

Key:

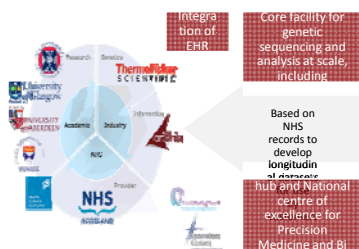
- 8.0 Green - achievable in a large region like Scotland or East of England
- 40.0 Amber - achievable across the whole of the UK
- White - requires international cooperation

Courtesy of Piers Mahon, Cancer Commons

Stratified Medicine Scotland
Innovation Centre

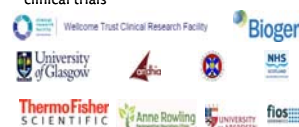
A New Approach To Collaboration. A platform for multidisciplinary collaboration:
AnalytiXagility is the informatics platform that supports clinical trials,

- A public/private joint venture - £30M investment to date to accelerate the development of precision medicine diagnostics and services
- Adoption by Scotland's National Health Service
- Also used as a distribution hub for AnalytiXagility's biomedical research informatics services
- Collaborative precision medicine development programs in arthritis, ovarian cancer, MS



Clinical Informatics & Biomarker Research
for MS: FutureMS

Clinical decision support for early
 prescribing of Disease Modifying
 Therapies in MS plus improving levels
 of treatment with better prediction of
 disease progression
 Collaboration across all Scottish
 Neurodegenerative centres to establish
 new standards of care informed by
 APOEε4 delivering analytical workspace
 for integration of clinical, image and
 genomic data to facilitate stratification
 of patients and improve standards of
 care, plus a clinical research workspace
 with Biogen to enable future federated
 clinical trials



Pan-European Alzheimer's Platform: EPAD (IMI and EFPIA Funded)

Evergreen cohort feeding proof of concept trials for prevention of Alzheimer's Disease

€64m project involving 36 EU partners.

Analytical workspace for balancing recruitment/cohort effectiveness, co-developing the EPAD engine for adaptive clinical trials



Challenges and Opportunities Towards a EU wide ecosystem



Big Issues 1: Complex environment



Interoperability: to work across systems with no additional effort

Big Issue 2. This is a tidal wave of data...



1987

1.44MB Floppy Disk
Marketed

Computer Science: Where are we Now

A **gigabyte** : 1000 megabytes

20 GB : Complete works of Beethoven

A **terabyte** : 1000 gigabytes

As of 2014, Wikipedia stored about 7 TB of information.

An **exabyte** : 1000 petabytes

Global Internet data: ~80 EB per month

A **petabyte** : 1000 terabytes

BBC iPlayer transfers 8 PB of programs every month.

1,000,000,000,000,000,000

A **zettabyte** : 1000 exabytes

World Wide Web: in 2015, holds 5 ZB of data...

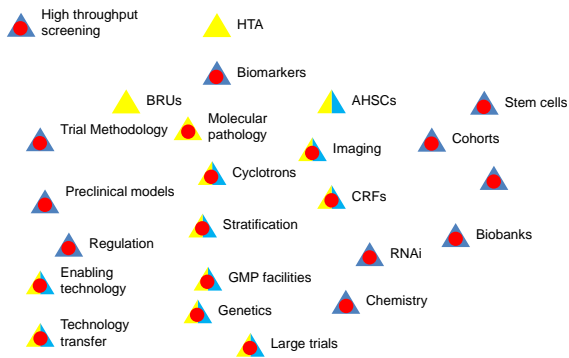


Health care is becoming increasingly data intensive



- **Internet and cloud** provide connectivity to every corner of the globe
- **Smartphone**: 2 billion users; 80% of adult population by 2020
- **Socialome**: the digital data harvested for health and wellness
- **Quantified Self**: Non-invasive biometric sensing, Tricorder wearables. Apple Research Kit
- **Exposome**: Pervasive environmental sensing will bring new knowledge to public policy decisions about creating a healthier physical environment, and
- **\$1000 genome** (genome, microbiome, transcriptome, lipidome, proteome, metabolome, multiome/panarome) Stem Cell and Genetic Tx (2000 + trials)
- **EHR data**: exponential growth of phenome from Electronic Health Records
- **Predictive Analytics** (Machine learning, AI, and Visualisation). Prediction: TenX more new knowledge from research in silico over RCT by 2020
- **Persuasive Technologies**: Behavioural and motivational sciences

Summary - the complex data intensive environment of medicine



Big Issue 3 : Direct-to-patient

23andMe Turns Spit Into Dollars in Deal With Pfizer

Direct-to-patient trial recruitment

ResearchKit: a clever tool to gather clinical data

"After six hours, we have 7,406 people enrolled in our Parkinson's study. Largest one ever before was 1,700 people" (Tweet from Parkinsons app developer, May 2015)

New software from Apple makes it easy for iPhone users to contribute to medical research but scientists warn there may be pitfalls.

Stanford Uni cardiovascular trial attracted 11,000 volunteers in one day after releasing their 'MyHeart Counts' App in the App Store. It would normally take a year to enroll that many subjects (Bloomberg Business, March 2015)

Early days. Sampling challenges (noise, representativeness), but great potential.

You're already carrying a powerful medical research tool.

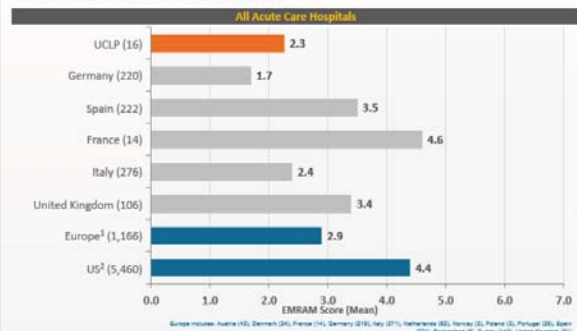
ResearchKit

Big Issue 4 Digital Maturity of Health Systems and Data

Level 8 - Personalized Medicine & Prescriptive Analytics
Level 7 - Clinical Risk Intervention & Predictive Analytics
Level 6 - Population Health Management and Suggestive Analytics
Level 5 - Waste & Care Variability Reduction
Level 4 - Automated External Reporting
Level 3 - Automated Internal Reporting
Level 2 - Standardized Vocabulary & Patient Registries
Level 1 - Enterprise Data Warehouse
Level 0 - Fragmented Point Solutions

EMR Adoption: World-Wide Benchmarks

HIMSS EMR Adoption Model Scores, Means per Segment, Q4/2014
(based on data from the last 24 months, no weighting applied)



Big Issue 5: Data Quality Data maturity and standardisation Pharmacogenetics of Metformin



ake one 2 twice a day for 1

1: 6: take onetab 3 times daily 1
2: 7: take one twice a day for 2 1
3: 8: take one at 8 am and 1 at 1
4: 9: three daily as directed 1
5: 10: two daily in addition to c 1
6: 11: 1 tablet eve meal 1
7: 12: 1 tablet five times dai 1
8: 13: take one twics a daily 1
9: 14: take one twice a daily 1
10: 15: 2 caplets twice a day 1
11: 16: take one 2 times/day with m 1
12: 17: 2 tablet 3 times daily cpus 4 1
13: 18: 1 tablet daily for 1 week then 1 tablet bd 1
14: 19: one 3 times daily 1
15: 20: one 5 times daily 1
16: 2 bd 116
17: 1 tablet bd 114
18: 1 tablet in the morning 103

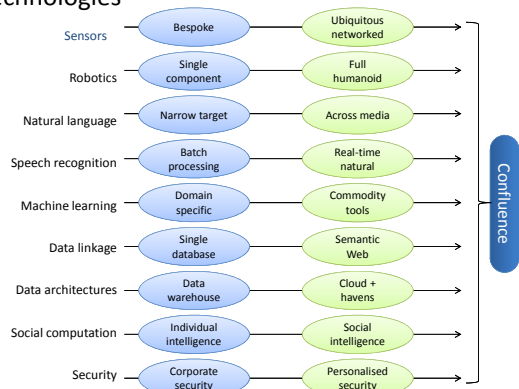


**5720 variations
for Metformin
alone!**

Big Issue 6 Harnessing Inter-disciplinarity



Technologies




Big Issue 7 (The biggest of all)



Public attitudes to the use and sharing of their data
Research for the Royal Statistical Society by Ipsos MORI
July 2014

1. The trust in data deficit

- Most institutions are trusted less with data than they are generally
- Younger people are more trusting with data than older
- Online retailers, internet companies, supermarkets, charities and academics have the biggest gaps between overall trust and trust with data
- Main concerns are usage creep, no personal benefit, and loss of data



Governance of Trustworthy Use of Data



- Increasing transparency & reducing uncertainty
- Agreeing standards: Principles & Best Practices
- Clarifying Responsibilities: Data Flows & Data Controllers
- Seeking buy-in from stakeholders

#datasaveslives schools engagement

Aims:

1. To inform and educate young people about the positive way that data can be used to improve public health
2. To inspire the next generation to consider careers in health data science





"100 Ways of using Data to make lives better"

- '100 Ways' will be a series of 100 national case studies which will showcase the leading health informatics research from the UK
- Aim: to raise the profile of the Institute and raise awareness of the benefits of using data in research
- Distribution will include sharing on social media under the #datasaveslives campaign.



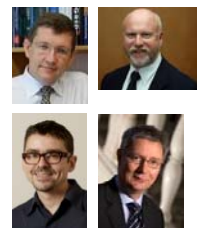
Federated Network of Safe havens "Diameter of Trust"

Safe Havens:

- ☐ One in each of the four NHS Research Scotland (NRS) nodes
- ☐ National Safe Haven
- ☐ Data enclaves – queries through remote access
- Working to common security and data sharing principles and standard
- Work collaboratively – allows data sharing, inter-operability
- External accreditation to provide assurance to data controllers, patients and the public



Opportunity to Scale across the nation The Farr Institute



Part of £200M investment

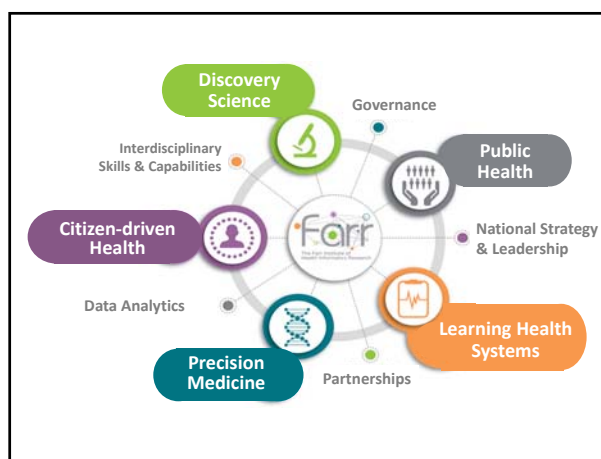


Our Six Key Activities



1. Cutting Edge Research
2. Harmonised eInfrastructure, methods, data curation
3. Public engagement.
4. Governance (safe havens)
5. Capacity Building
6. Partnerships

To deliver impact nationally and internationally





EPCC – national service provider



- EPCC is the UK's national HPC provider
- ARCHER and RDF - £96m UK Gov investment
- 3,500 users
- 118,080 cores
- 28Pb of disk storage
- Managed alongside Farr Institute system at ACF
- Data Security

- Physical sciences have dominated HPC provision for 20 years
- Limited use by biosciences and medicine
- Technology is bringing HPC and Data Analytics together
- New users from medicine and genomics dominant
- 600 projects



Options and Opportunities for Health Data Science

- **Commitment** to communication, collaboration and public engagement in everything we do
- **Convergence** of care with innovation and research
- **Clinical data quality** – Bringing routinely collected data up to same standard as high quality research data
- **Collaboration** of health care providers/ academia
- **Commercial engagement** - encouraged but with transparent governance, collaborative and benefit sharing
- **Computer Science** key ingredients for change – educational implications
- **Clarity** about GOVERNANCE and data sharing

An Opportunity for EU to Lead the Way?

This is non-negotiable!

It is not the strongest of the species that survive,
nor the most intelligent,
but the one most responsive to change



Charles Darwin 1809-1882

Believe that you can
and you are half way there!



FD Roosevelt 1882-1945



This will require international cross-border working!

Thank you for listening