All of us together

UK Health Data Research Alliance Symposium

Tuesday 1 December 2020
09:30 - 17:00

Priority datasets to answer priority questions

Chair: Alison Pritchard, Office for National Statistics

HDRUK
Health Data Research UK

UK Health Data Research Alliance

Health Data Research Innovation Gateway

→ Join the conversation:
hdruk.ac.uk/HDRAlliance20
@HDR_UK
#HDRAlliance20

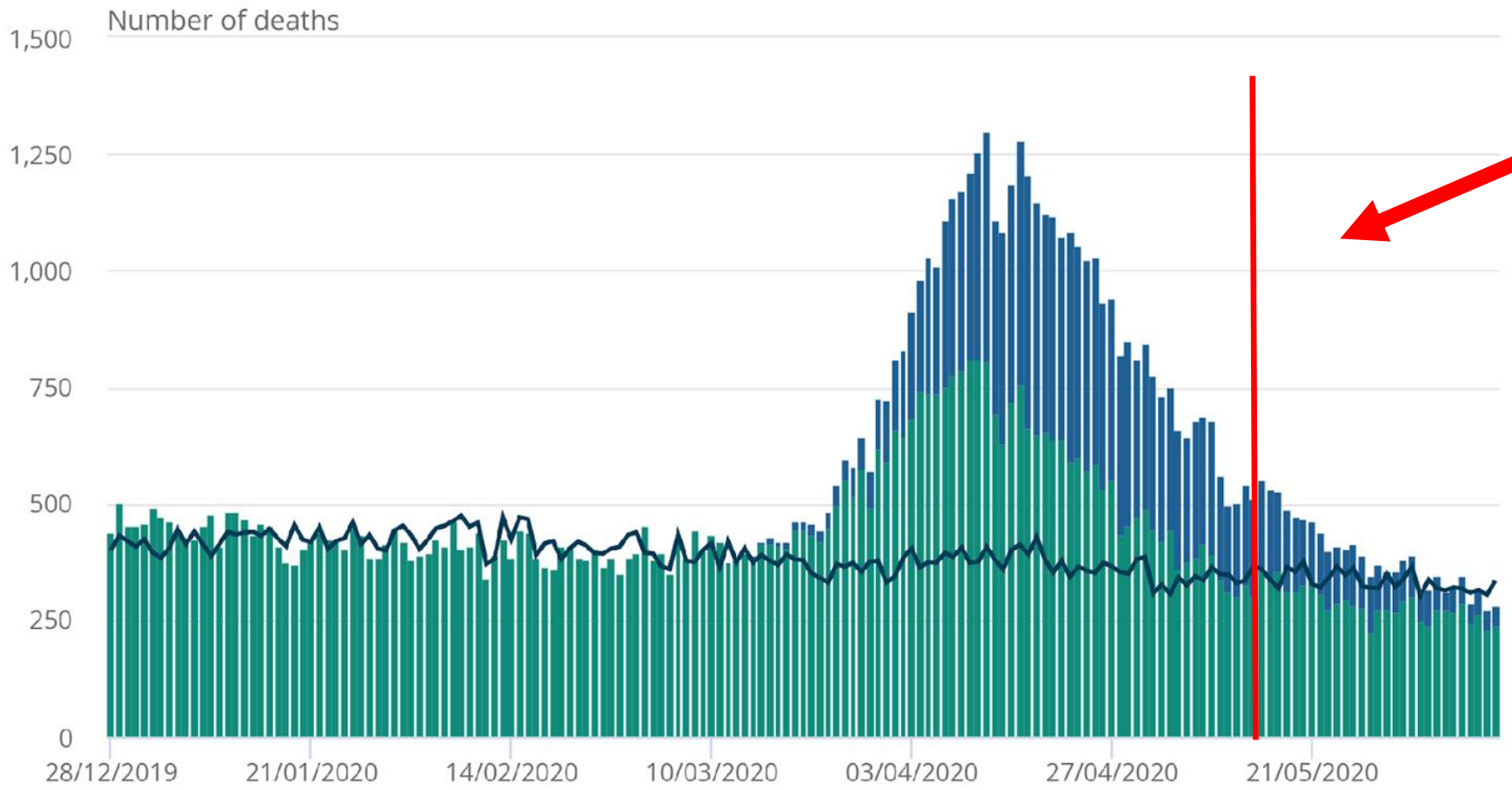
FACULTY OF POPULATION HEALTH, INSTITUTE OF HEALTH INFORMATICS

Estimating prevalence of SARS-CoV-2 in care homes: The VIVALDI-1 survey

Dr Laura Shallcross,
l.shallcross@ucl.ac.uk

vivaldi
Preventing COVID-19 in care homes

Number of deaths of care home residents from 28/12/19 – 12/6/20, England and Wales



Introduction of whole care home testing in England

SOURCE: ONS

Different approaches to measure prevalence

PILLAR 1 TESTS (PHE/SUS)

Hospitalised patients (NHS IDs)
PHE Outbreak investigations
Not linked to care homes (No CQC ID)

PILLAR 2 TESTS (NHS FOUNDRY)

All staff and residents
Variable uptake and irregular testing
CQC ID usually recorded
60% NHS ID's
Differentiates: Symptomatic/Asymptomatic
Staff versus resident

OUTBREAK NOTIFICATIONS (HPZone)

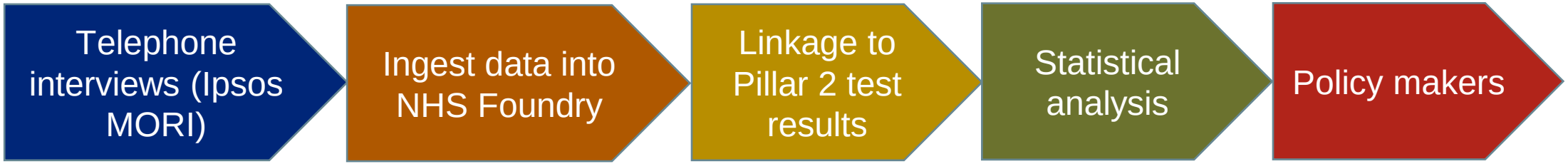
Suspected/confirmed cases
Variable testing
Under-reporting

SELF-REPORTED (SURVEY)

Care Home managers
No lab confirmation
Recall bias
Can capture data on care home characteristics
Ethical approval

VIVALDI-1 survey – study design

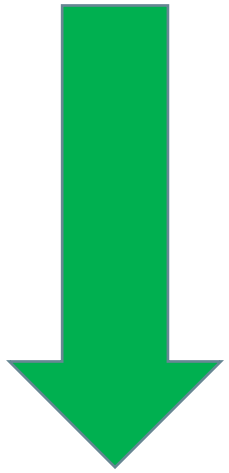
- Collaboration between UCL, ONS, DHSC and PHE
- Telephone survey of care home managers (26 May – 19 June)
- English care homes mainly providing dementia care or care to > 65 years
- Outcomes:
    - Self-reported confirmed infections*
    - SARS-CoV-2 test results (whole care home testing programme)
    - Risk factors for infection, outbreaks and large outbreaks#



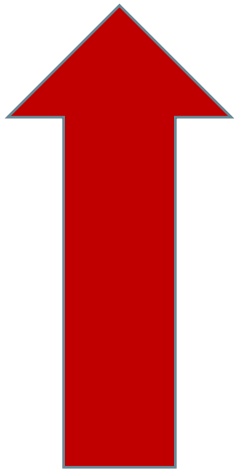
*Number of confirmed cases reported to the care home since the start of the pandemic as a proportion of the total ; # > one third of staff/residents infected or at least 20 cases per care home

Weighted prevalence of and risk factors for SARS-CoV-2 infection (n=5126 care homes)

Participants	Survey (95% CI)*	Pillar 2 testing (95% CI)
Resident n=160,033	10.5 % (9.9-11.1%)	2.8% (2.4-3.1%)
Staff n =248,594	3.8% (3.4-4.2%)	0.6% (0.5-0.8%)



REDUCE RISK
Staff sickness pay
Cohorting staff
No use of agency (temporary) staff



INCREASE RISK
Difficulty isolating residents
For profit care homes
More admissions to the care home

*Proportion of staff and residents with confirmed positive test since 1st March 2020. Pre-print available from https://www.medrxiv.org/content/10.1101/2020.10.02.20205591v1

Acknowledgements

- **UCL**
- Maria Krutikov
- Chris Fuller
- Andrew Hayward
- Andrew Copas

- **ONS**
- Danielle Burke
- Sapphira Thorne
- Owen Abbott
- Katie Sharp
- Leone Wardman

- **DHSC**
- Alasdair Donaldson
- John Hatwell
- Jane Cummings

- **Other**
- Susan Hopkins (PHE)
- Gemma Hallatt (Palantir)
- Ipsos MORI

Pre-print available from:
https://www.medrxiv.org/content/10.1101/2020.10.02.20205591v1

HDR UK
Health Data Research UK

mprov ng our health
through data sc ence

Using linked data to respond to the COVID-19 pandemic in Wales

8th October 2020

Dr Rich Fry

on behalf of the SAIL COVID Team

Swansea University Prifysgol Abertawe | Medical School Ysgol Feddygaeth

SAIL DATABANK

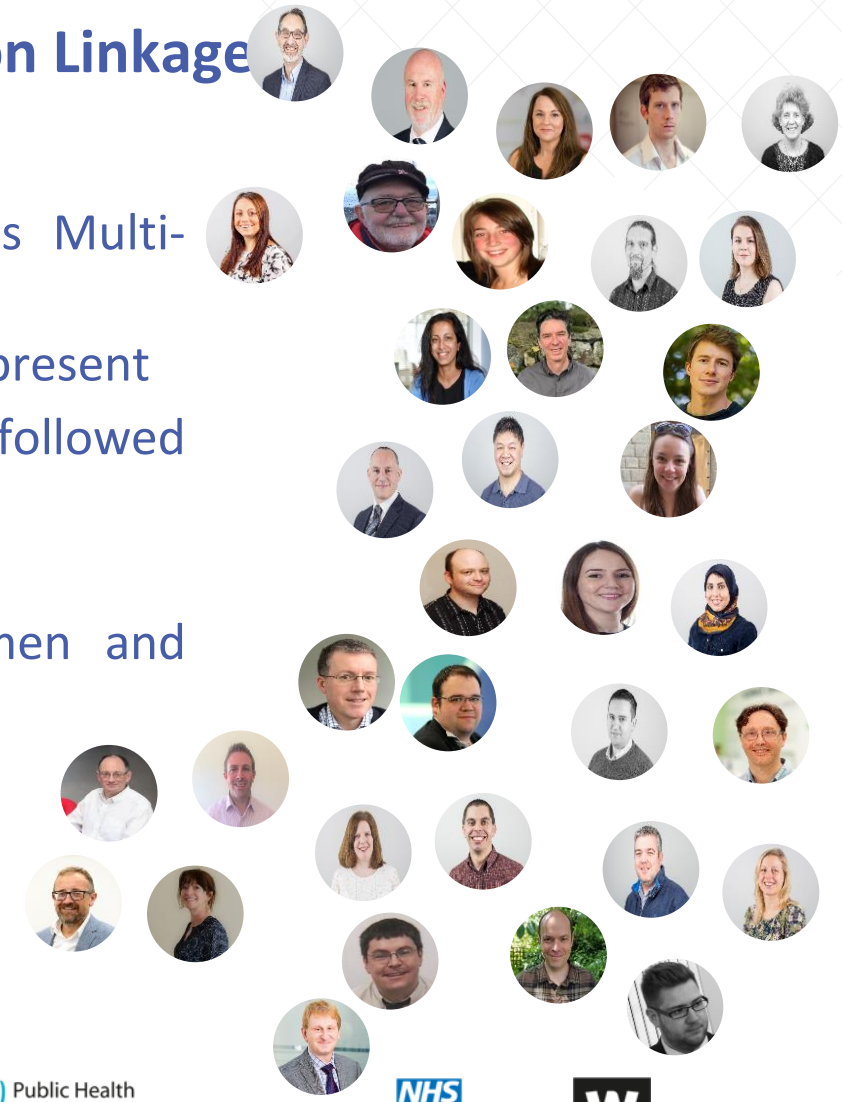
ADR WALES Data-driven change

Population Data Science at Swansea University

Analysis of linked de-identified data of the impact of COVID19 on the Welsh population using the Secure Anonymised Information Linkage (SAIL) system: A One Wales Approach

Creation of two total population linked cohorts derived from Wales Multi-morbidity Cohort:

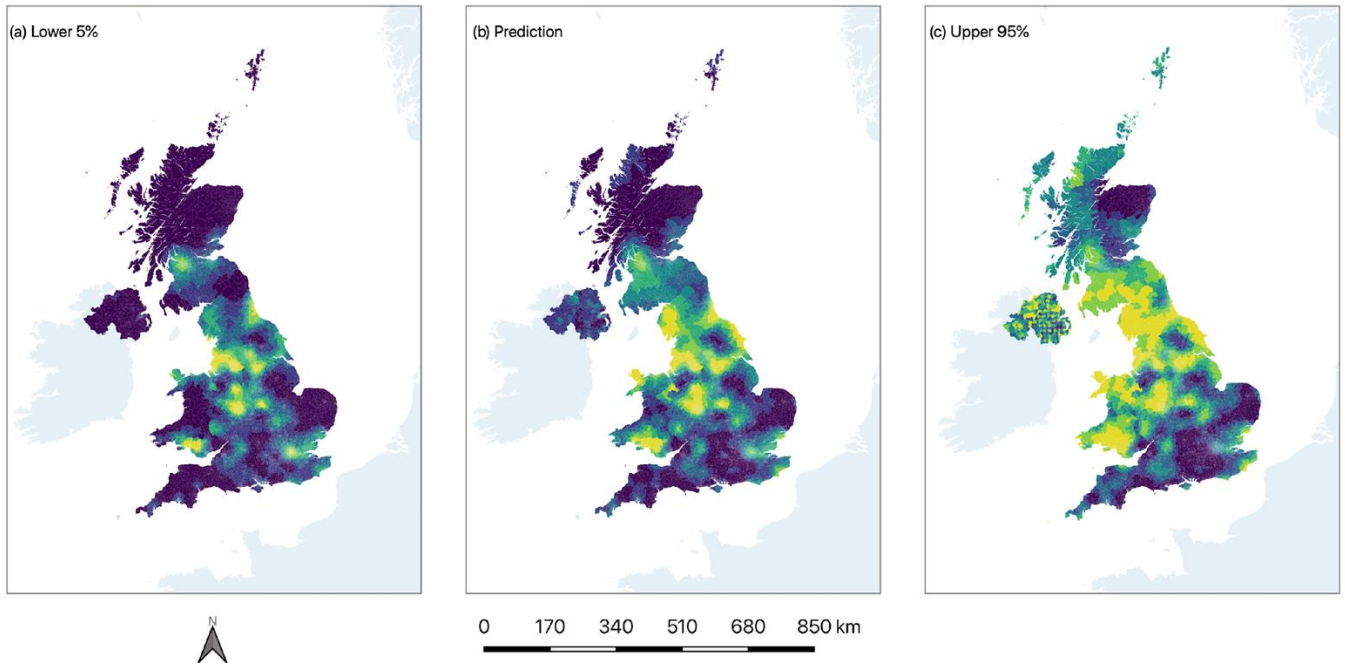
- C20 - all alive and known to NWIS on 1st January 2020 followed up to present
- C16 – all alive and known to NHS Wales (NWIS) on 1st January 2016 followed up to end December 2019 – counterfactual cohort
- Research now supported by a grant from the Medical Research Council
- Enhanced data collection in care homes, schools, pregnant women and teenagers
- Advanced Spatial Analytics
- Direct reporting to Welsh Government TAG and to SAGE
- https://bmjopen.bmj.com/content/10/10/e043010.info

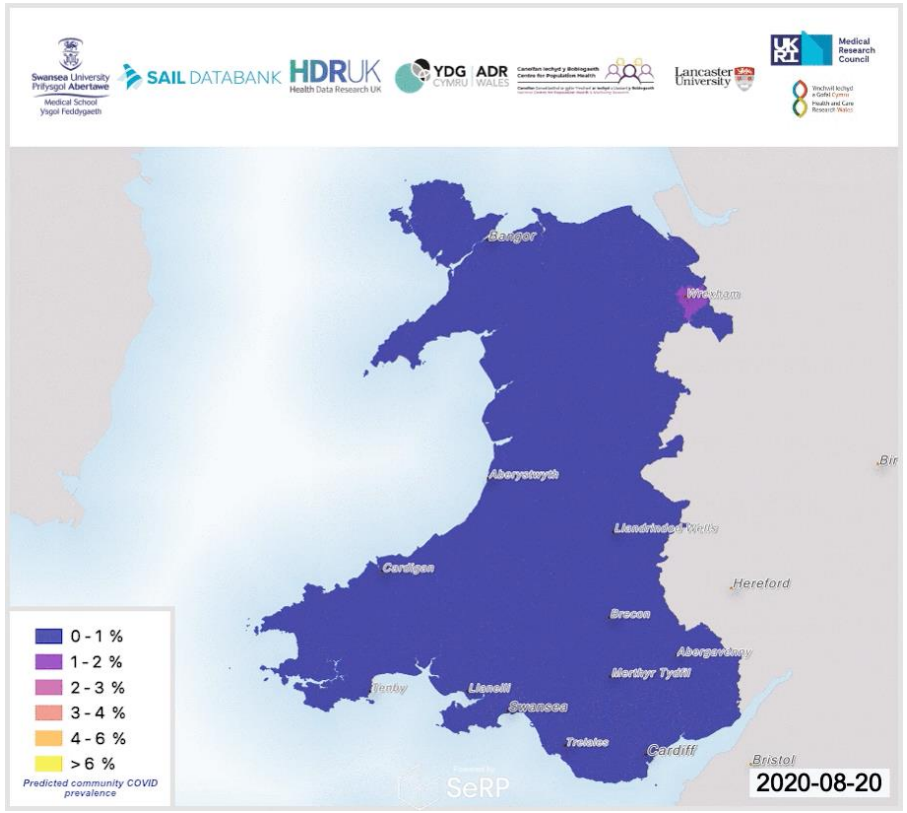


Predicting community prevalence

- HDRUK Network Linkup with BREATHE, ZOE and geospatial expertise to develop high-resolution mapping of prevalence across the UK. April Present
  https://www.medrxiv.org/content/10.1101/2020.08.17.20175117v1

- First UK wide scale mapping of prevalence at community level – before testing data was available. Weekly updates to Welsh and Scottish Gov. Importantly shows intra-authority variation
- Models adapted and refined to use testing data and other covariates – used as part of the evidence by First Minister for October/November national lockdown

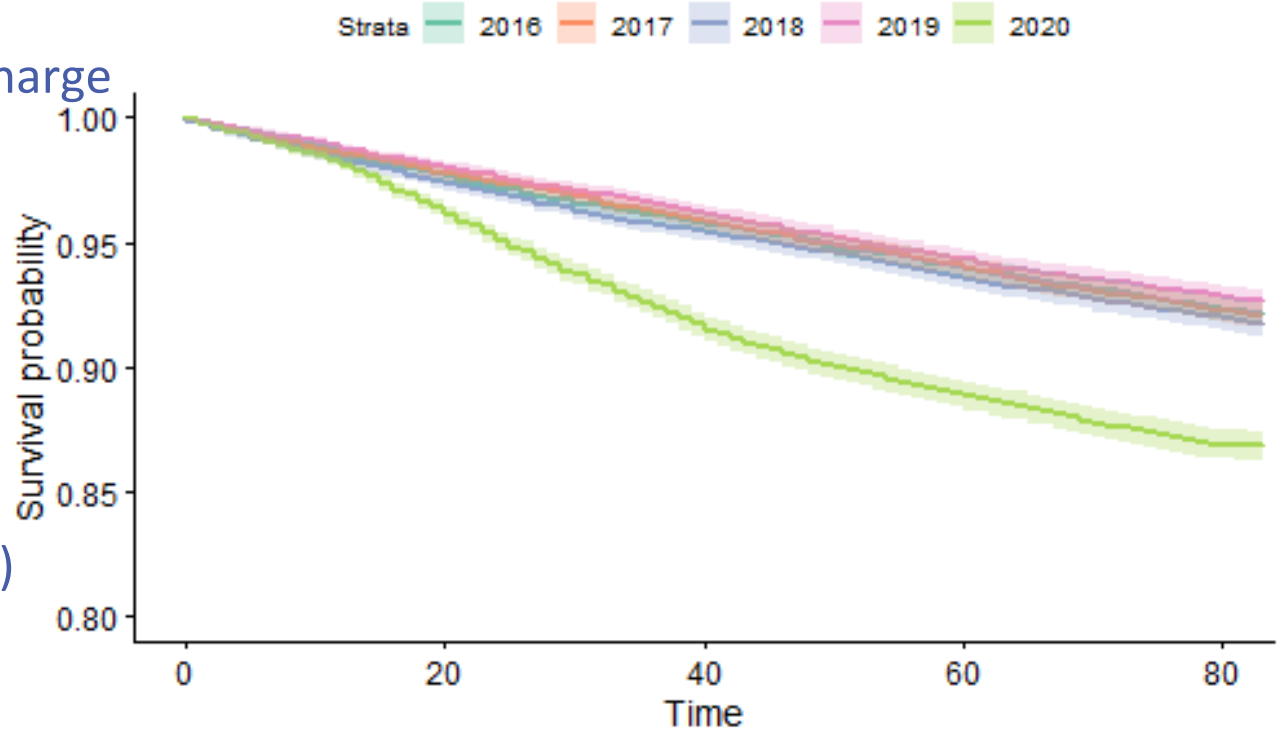




Care Homes

TAG, Welsh Government Task and Finish Group & SAGE SCWG

- Enhanced Care Home Index – total care home linkage plus enhanced variables to capture lived environments (e.g. Floor space, linked care homes, shared space, services offered, access to primary and secondary care)
- Initial analysis: excess mortality (all cause)
- Discharge analyses : showed small effect of discharge



Ongoing work:

- Looking at COVID specific mortality
- Change in resident population characteristics
- Impact of community prevalence on outbreaks
- Impact on specialist care settings (e.g. dementia)
- Care home workforce analysis

https://doi.org/10.1093/ageing/afaa207

TEAM SCIENCE!

Powered by

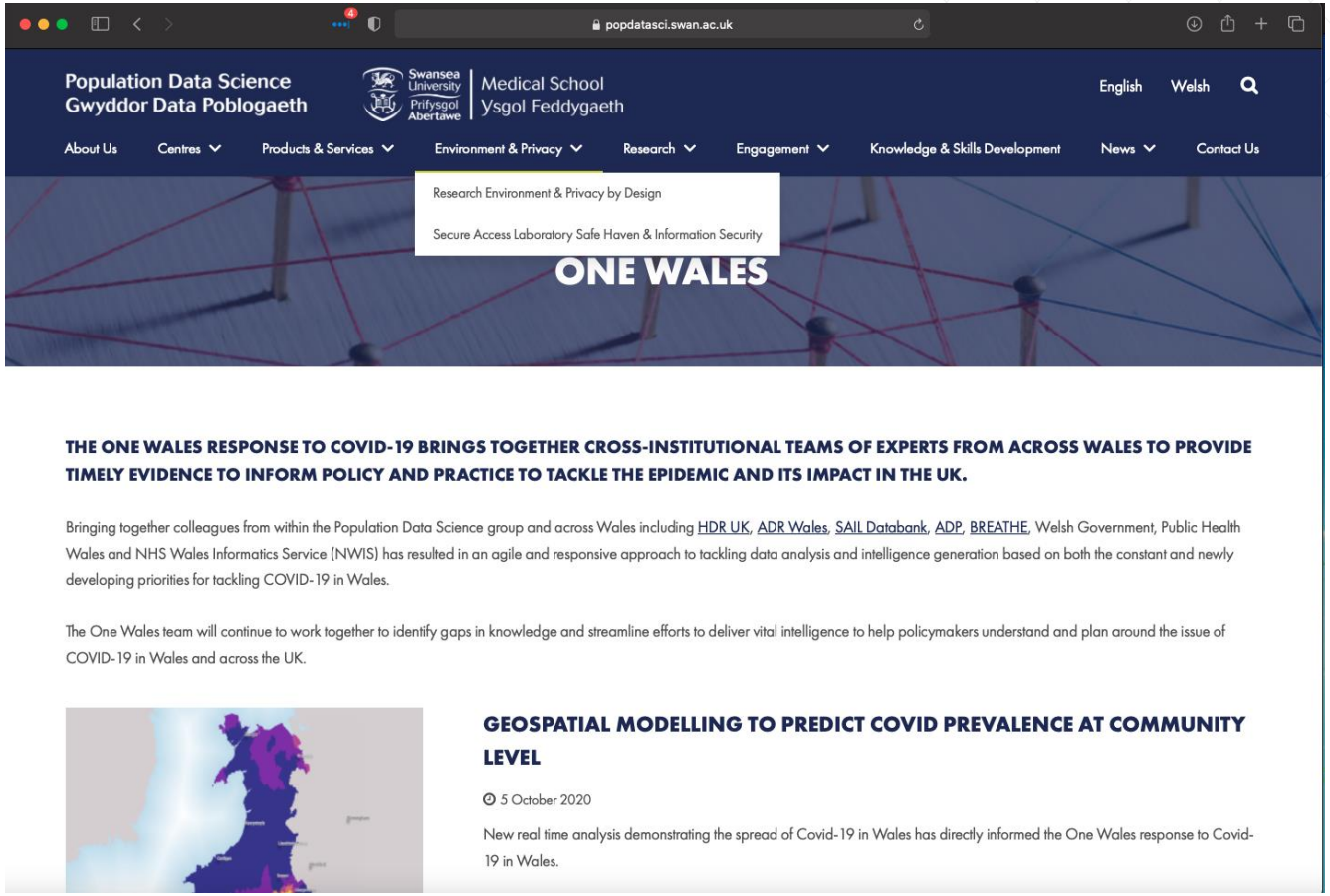
SeRP
ISO certified and UK Statistics Authority Accredited

Thanks for listening!

Email: r.j.fry@swansea.ac.uk

@richfry

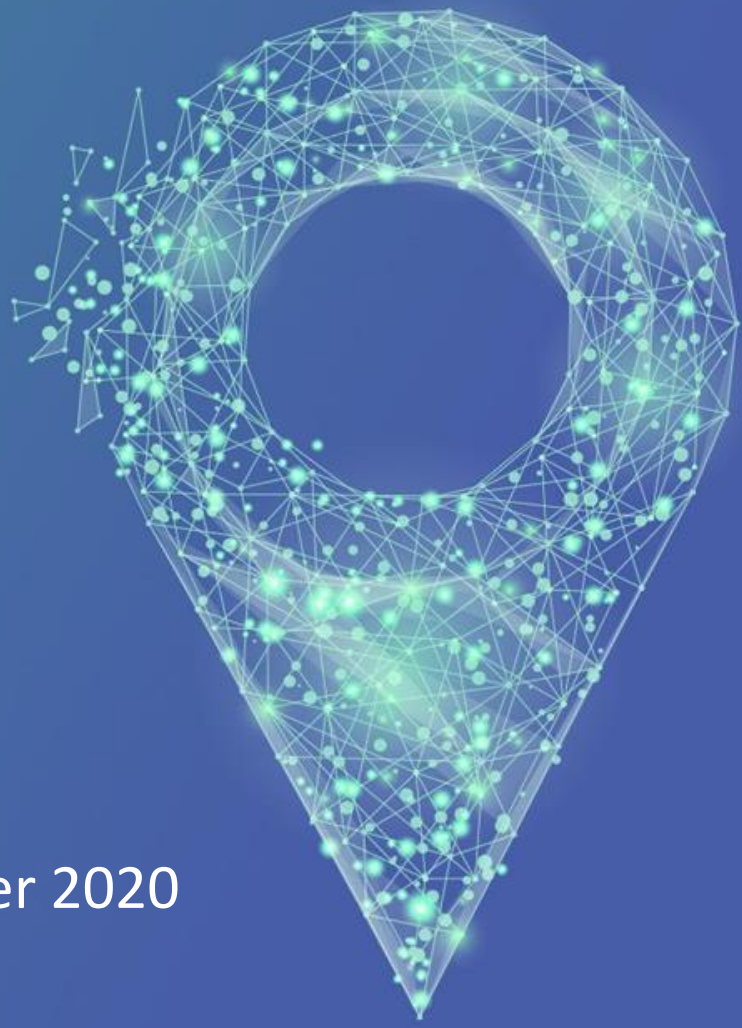
HDRUK
Health Data Research UK



https://popdatasci.swan.ac.uk/news/one-wales/

HDRUK
Health Data Research UK

British Heart Foundation
Data Science Centre
Led by Health Data Research UK



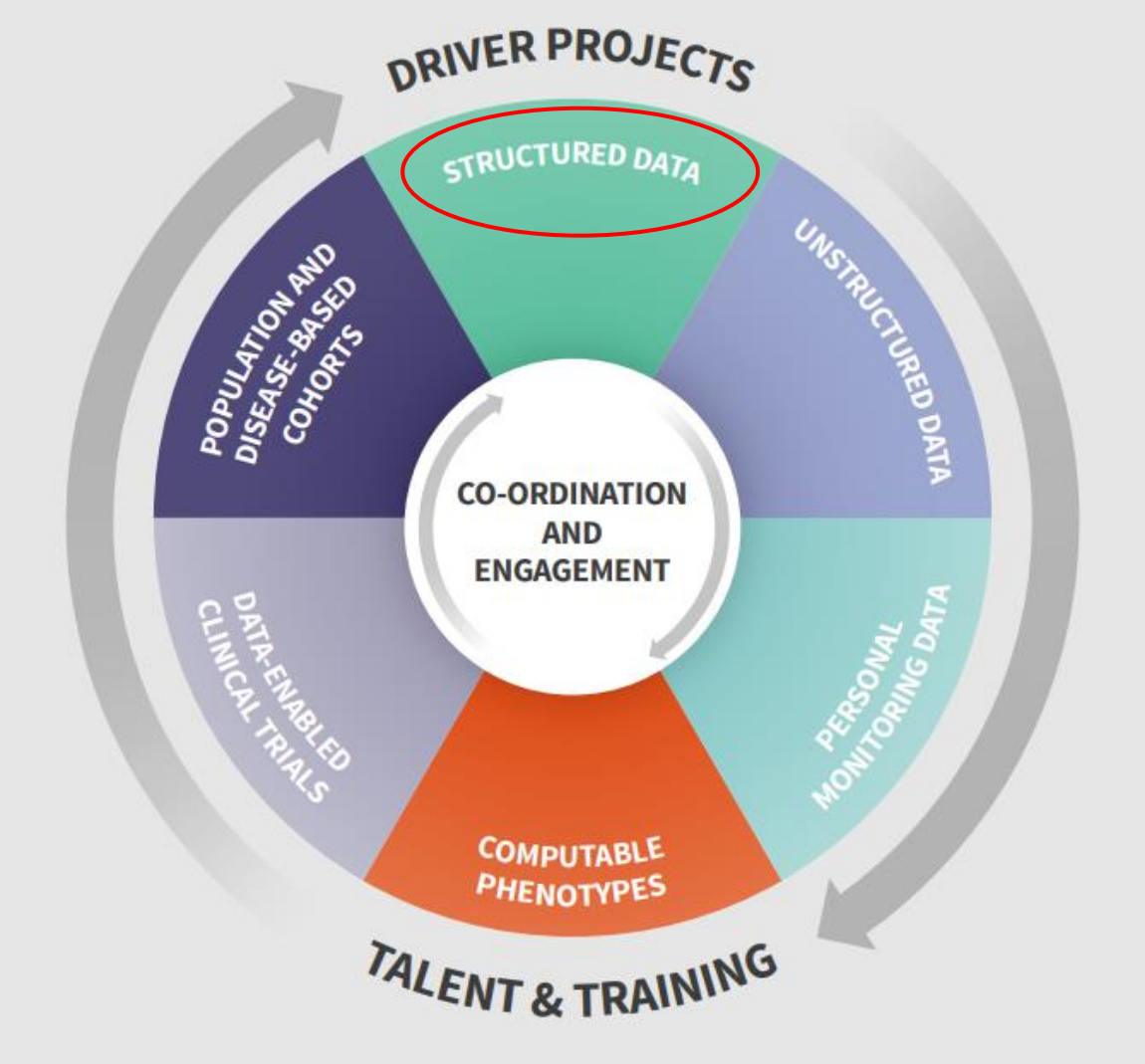
UK Health Data Research Alliance Symposium, December 2020

Priority datasets for priority questions

Our strategy and delivery plan

British Heart Foundation Data Science Centre
Led by Health Data Research UK

HDRUK
Health Data Research UK



6 thematic work streams:

- Structured, coded health data UK population-wide
- Unstructured health data
- Personal monitoring data
- Computable cardiovascular phenotypes
- Data-enabled cardiovascular trials
- Population and disease-based cohorts

and

3 cross-cutting work programmes:

- Co-ordination and engagement
- Driver projects
- Talent and training

COVID-19 and cardiovascular disease: the CVD-COVID-UK consortium driver project

British Heart Foundation Data Science Centre
Led by Health Data Research UK

HDRUK
Health Data Research UK

UK-wide consortium of universities, NHS bodies and data custodians

Linking routinely collected data from across the whole population of the UK to address three critical questions:

1. What is the impact of cardiovascular disease (and its risk factors and medications) on susceptibility to and outcomes of COVID-19?
2. What is the direct impact of COVID-19 on cardiovascular diseases?
3. What is the indirect impact (unintended consequences) of COVID-19 on cardiovascular diseases?

CVD-COVID-UK: progress

British Heart Foundation Data Science Centre
Led by Health Data Research UK

HDRUK
Health Data Research UK

- Ethics approval in place (May)…regulatory approvals followed
- Inclusive consortium grown to include around 100 members
- Consortium principles developed
- NHS Digital's first-ever TRE now up and running to support this programme (from mid-July)
  - pn-boarding of datasets ongoing – slower than hoped for but moving
  - > 20 analysts from > 7 institutions accessing data; core group of highly active analysts
  - regular organisational and user/technical meetings with NHS Digital, unearthing and solving issues iteratively (long game)
  - more analysts and institutions being added
  - 5 active projects (medications, direct and indirect impacts on CV disease, methods), initial outputs
  - project applications process in place to enable other projects in advanced stages of planning
- Access to similar data in Scotland and Wales now approved and access being set up
- Approvals and Oversight Board established and first meeting held
- NHS Digital data wrangler service set up
- Great feedback from research community, patients and public partners – transformational

CVD-COVID-UK: building UK-wide infrastructure to accelerate UK-wide research

British Heart Foundation Data Science Centre
Led by Health Data Research UK

HDRUK
Health Data Research UK

Public Health England
- **COVID 19 lab test data**
- ~~**CHESS dataset**~~

Intensive Care National Audit & Research Centre
- Intensive care dataset

NHS Digital
- **Hospital data (HESx4 + SUS)**
- **Death registry data**
- **Primary care data**
- **Community dispensing data**

Cardiovascular data
- Cardiology, cardiac surgery, coronary intervention audits (NICOR)
- Stroke audit (SSNAP)
- National vascular registry

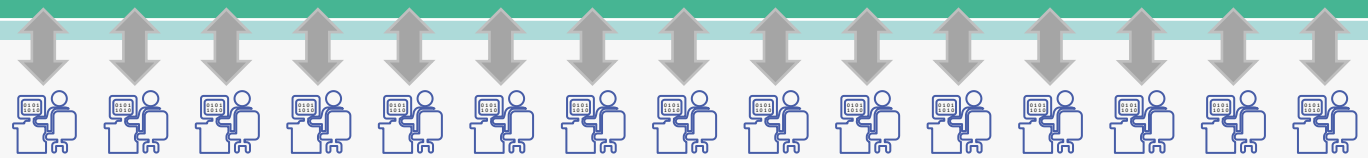
NHS Digital link datasets using NHS number

NHS Digital Cardiovascular Trusted Research Environment

Secure storage and de-identification of linked datasets

Secure access for approved researchers

CVD COVID UK de-identified linked dataset



Approved researchers

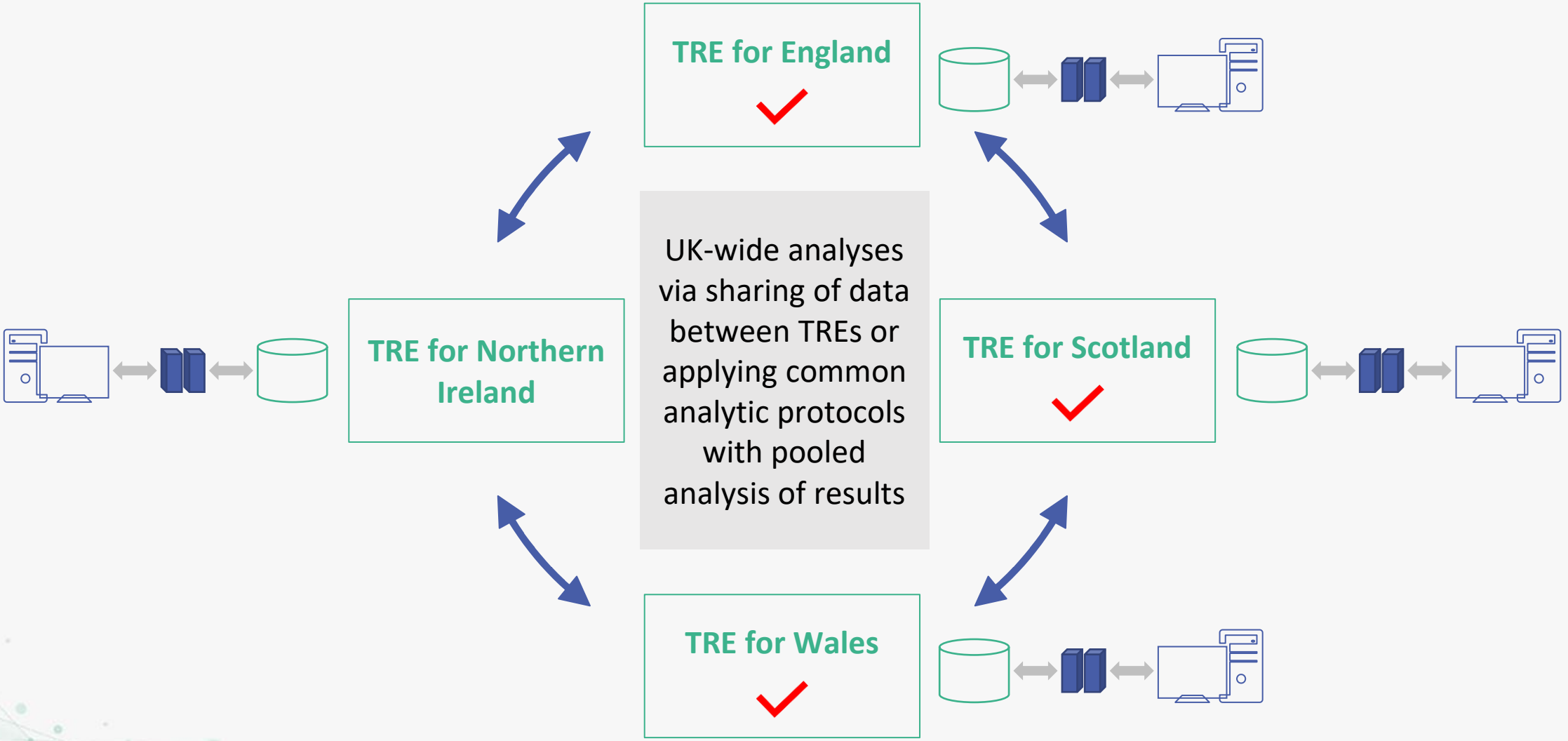
Secure data flows

Secure access to de-identified data and aggregate outputs

CVD-COVID-UK: building UK-wide infrastructure to accelerate UK-wide research

British Heart Foundation Data Science Centre
Led by Health Data Research UK

HDRUK
Health Data Research UK



CVD-COVID-UK: building UK-wide infrastructure to accelerate UK-wide research

British Heart Foundation Data Science Centre
Led by Health Data Research UK

HDRUK
Health Data Research UK

Data type	Country			
	England (NHS Digital TRE)	Scotland (National Data Safe Haven)	Wales (SAIL)	Northern Ireland (Honest Broker Service)
Population size (approx)	**58 million**	**5.5 million**	**3.2 million**	**1.6 million**
Hospital data (e.g. HES, SMR, PEDW)	Available	Nov 2020	Nov 2020	TBC
COVID lab testing data (e.g. SGSS, ECOSS)	Available	Nov2020	Nov 2020	TBC
Primary care data	Available	Nov 2020	Nov 2020	TBC
Death data	Available	Nov 2020	Nov 2020	TBC
ITU data	TBC	Nov 2020	TBC	TBC
ITU/HDU admissions (CHESS)	Available	N/A	N/A	N/A
Prescribing/Dispensing data	Available	Nov 2020	Nov 2020	TBC
NICOR CVD data	Nov 2020	N/A	Dec 2020	TBC
Stroke audit data	Nov 2020	Dec 2020	Dec 2020	TBC
National Vascular registry	Nov 2020	N/A	TBC	N/A

GP Data Exploration

Jenny Cooper

10 September 2020

GP Data (gdppr_dars_nic_391419_j3w9t)

Overview

The GP Data provides a line per clinical recording for a patient. There are therefore multiple rows per patient. The record_date ranges from 1793-12-07 to 2919-12-18. The plot below shows completeness of data from 2015 to present.

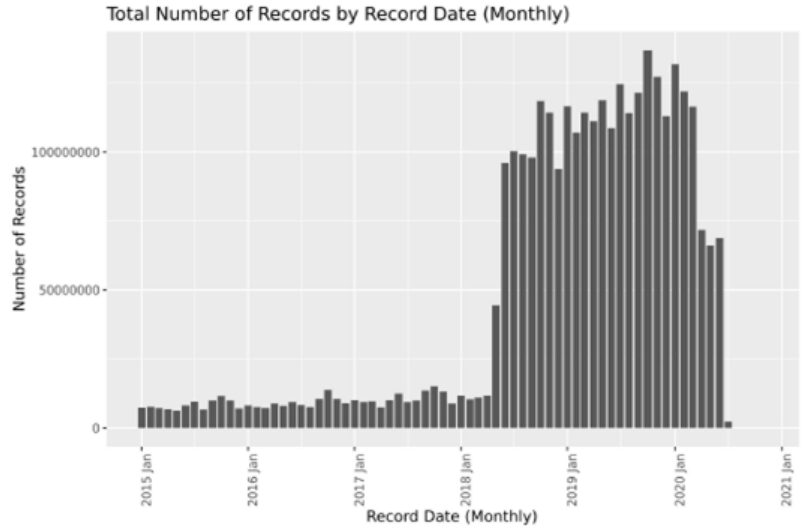


Figure 1: Total number of records using Record Date from 2015 to present (monthly)

4 billion records
56.5 million patients