

# Link between clinical coding /data quality, & population health

## *Learning from the Oxford-RCGP Research & Surveillance Centre (RSC)*

Simon de Lusignan

@Lusignan\_S

GP & Professor of Primary Care &  
Clinical Informatics

Director RCGP Research & Surveillance  
Centre (RSC)



## Simon de Lusignan

- **Practising GP**
  - >37 years in my practice
  - Microcosm of current pressures on primary care
  - 16,000 registered patients in the space for 8,000
  - New premises soon?
- **Professor of Primary Care & Clinical informatics**
  - Nuffield Department of Primary Care Health Science
  - Director Oxford-RCGP Research & Surveillance Centre (RSC)
  - Research interests: routine health data; technology enabled new roles in primary care, & self-management
  - ***Coding is caring! Sampling is informing!***
- Governing Body Fellow Green Templeton College



Royal College of  
General Practitioners  
Research & Surveillance Centre



# Summary

1. Population health is a health system objective
2. Data for surveillance - RSC sentinel network
3. Summary – Developing the primary care data custodian



# Population health is one of the five goals of health systems

## Improving population health is one of the five goals (quintuple aims) of health systems

Population health can't be discussed in isolation, it must be delivered in the context of:

- Patient experience
- Ensuring health equity
- Cost control
- Care team wellbeing

The focus of this presentation will be how routine data held on computerised medical record (CMR) systems can support population health... ..easier where data are coded & of good quality. Coding is caring!!



### References:

DOI: [10.1377/hlthaff.27.3.759](https://doi.org/10.1377/hlthaff.27.3.759)

doi: [10.1370/afm.1713](https://doi.org/10.1370/afm.1713)

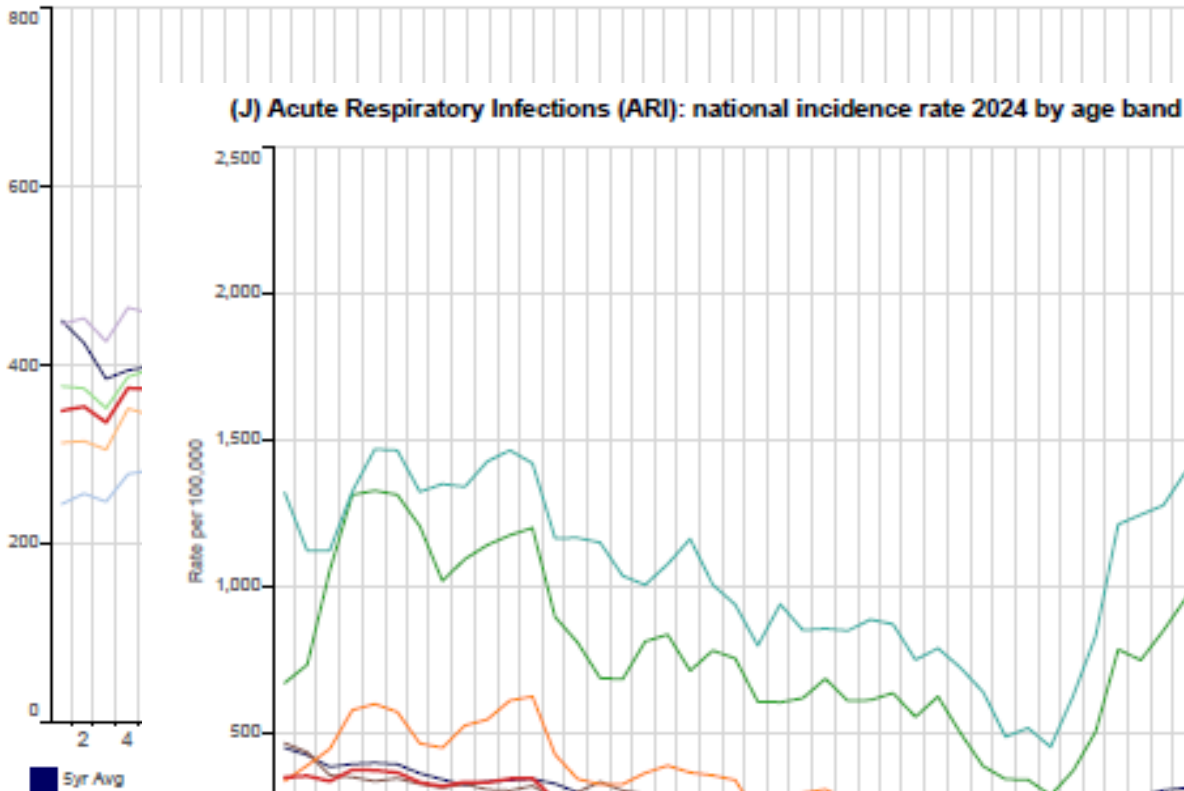
DOI: [10.1001/jama.2021.25181](https://doi.org/10.1001/jama.2021.25181)

DOI: [10.1136/bmj.n1262](https://doi.org/10.1136/bmj.n1262)

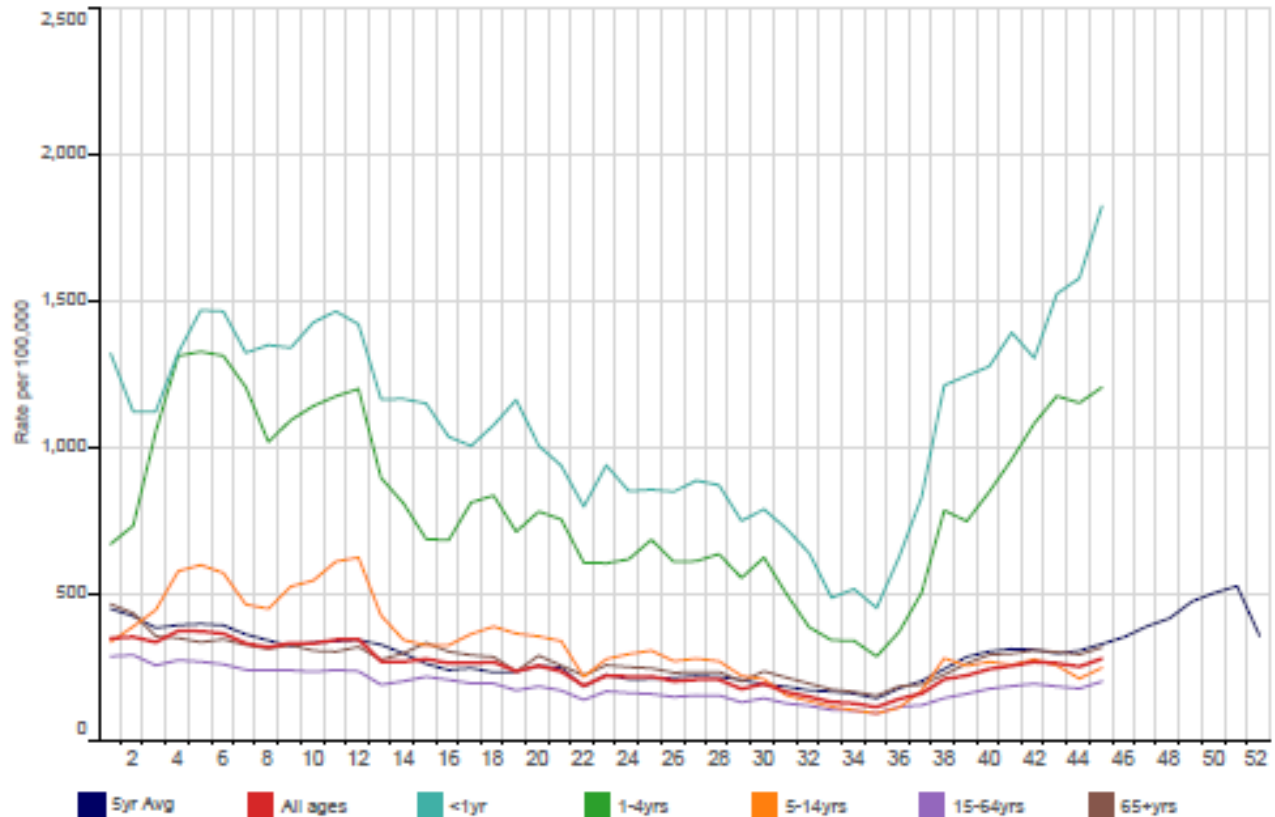
# The annual pattern of acute respiratory infections (ARI) ...

The annual pattern of respiratory infection (ARI) might help with planning...

(I) Acute Respiratory Infections (ARI): national incidence rate 2024 by region



(J) Acute Respiratory Infections (ARI): national incidence rate 2024 by age band



# Focus on data (of quality) required for population health

## PICO can provide a framework for investigating population health

We will explore where we can use routine data to define:

**Population:** We have a registration-based system, but there is denominator inflation

**Interventions & exposures:** Can be recorded well

**Comparators:** Matching is a challenge

**Outcomes:** Many important outcomes are not recorded in primary care data

**P** = Population

**I/E** = Intervention  
/ Exposure

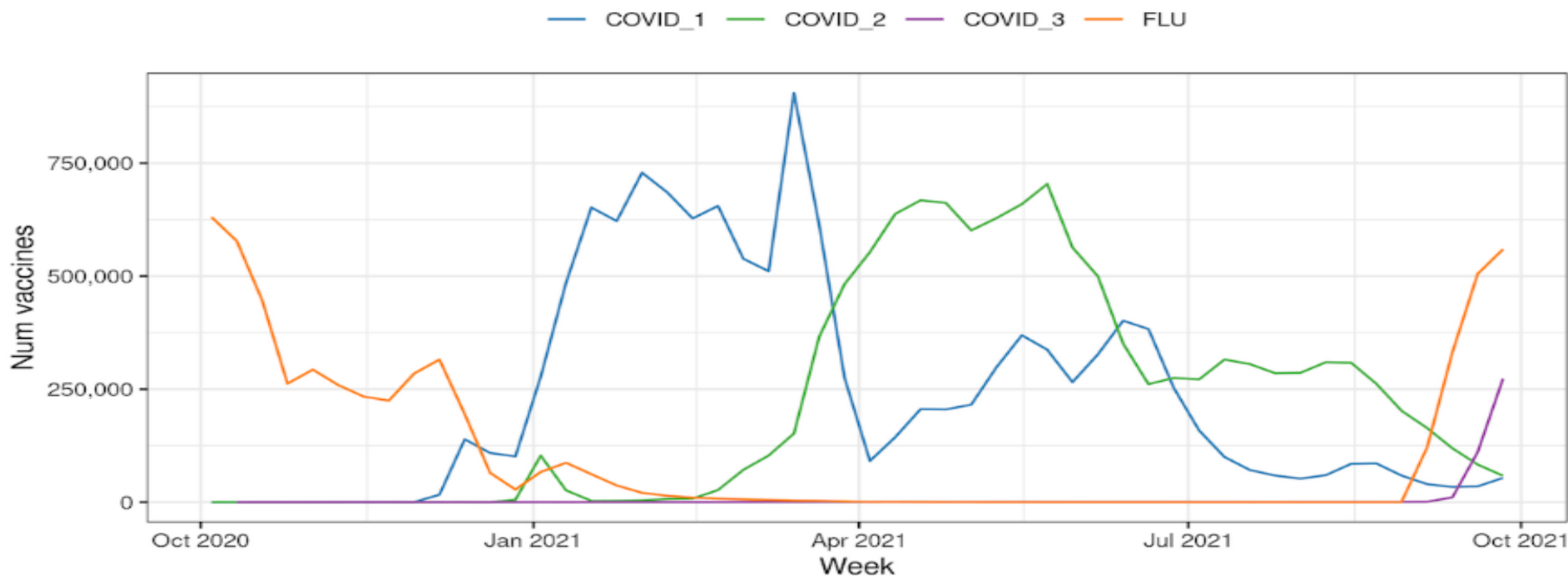
**C** = Comparator

**O** = Outcome

# COVID-19 & influenza vaccine exposure data

- UK primary care records capture vaccine exposure data
- Key to be combined with population and cases of COVID-19 and influenza for assessing vaccine effectiveness (VE)
- Data October 2020 to 2021

Reference:  
DOI: [10.2196/39141](https://doi.org/10.2196/39141)



## GP computerisation enables population health

**Early Computerisation (1970s-1980s):** - Minority of GPs. Benefits repeat prescribing & recall.

### **Government initiative (1990s-2000s)**

- **Reimbursement of computerisation costs:** Computers for general practice scheme (from 1990)
- **Pay-for-performance (P4P):** For chronic disease management, based on routine clinical data

### **NHS initiatives to integrate data**

- **NHS number:** unique health system wide identifier, enables linkage to key outcome data
- **Laboratory links:** enabling pathology to be made from the computerised medical record (CMR) with results returned automatically
- **Electronic prescribing:** All prescription requests and issue are done online (not UK-wide)

### **Development of specialist primary care applications (EMIS, TPP, others)**

Unique for the UK environment they are universally used . Standards set for suppliers

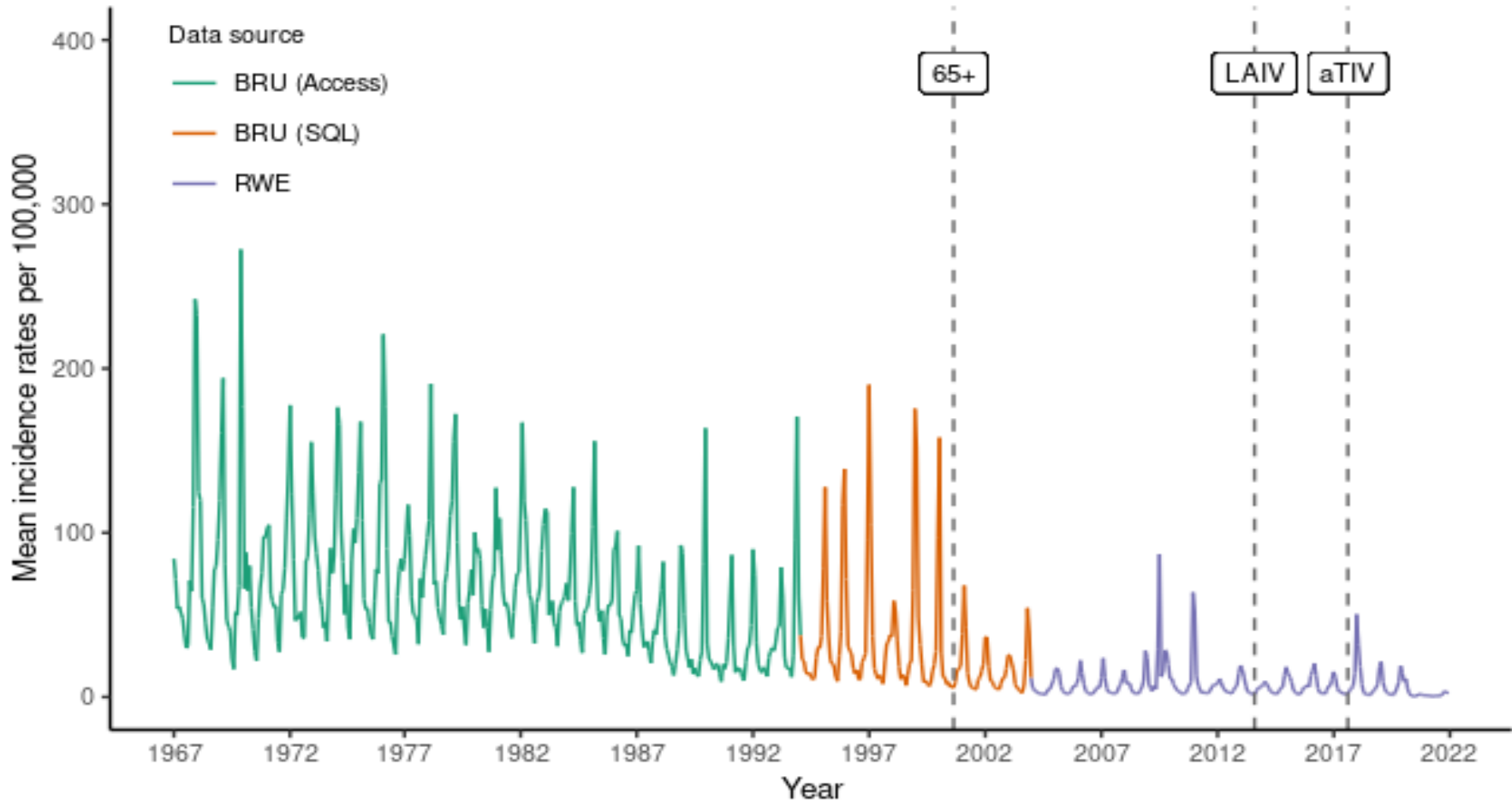
**Patient access to their records:** - Progressively increased patient access to records.

**Remote consultation:** - Accelerated by the pandemic, + same day “total triage”



# Primary care data unique longitudinal insights – ILI & flu vaccines

### Mean 4-weekly incidence rates of influenza-like illness



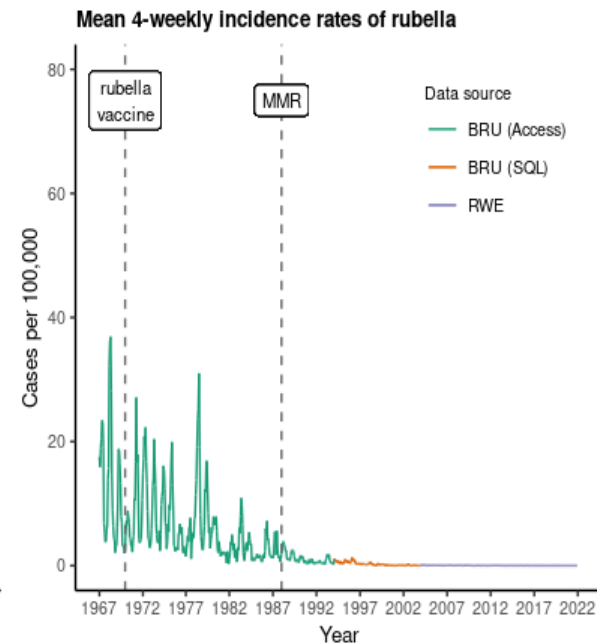
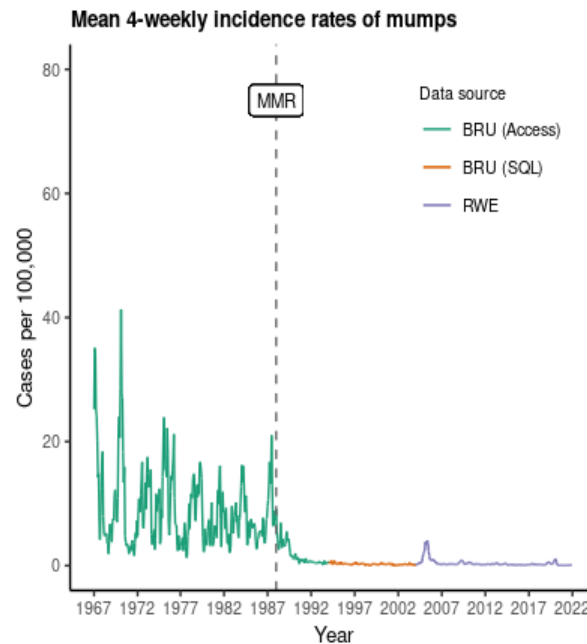
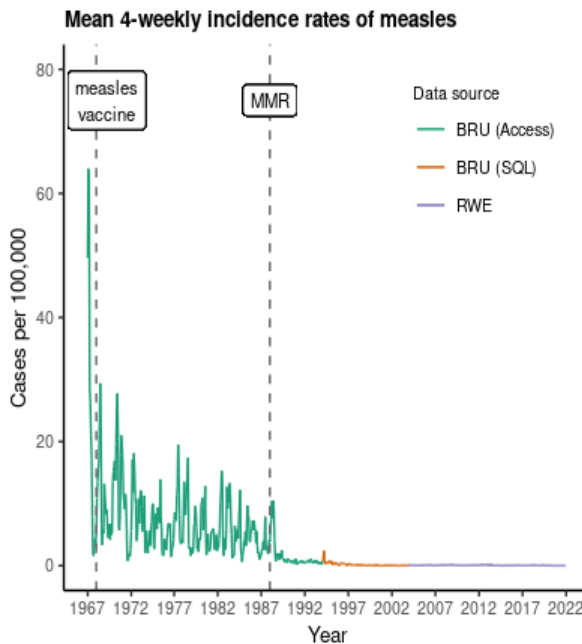
# Primary care data provided unique longitudinal insights - MMR

- Longitudinal data from the RSC – Wellcome funded project to link historic data
- Measles vaccine introduced in 1968, rubella in 1979 –initial decline in incidence
- Rates very low for all three conditions after introduction of MMR in 1988

## Measles

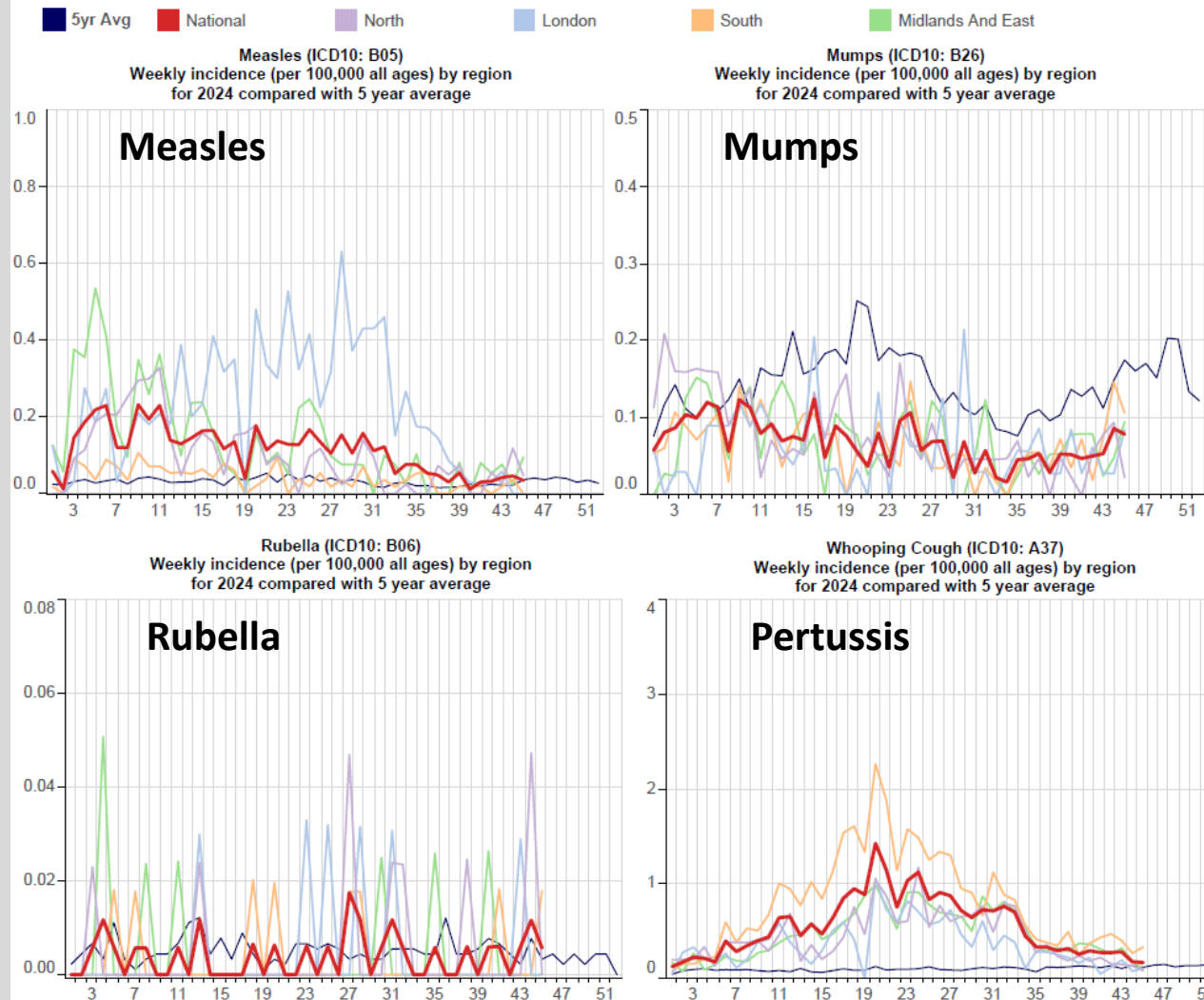
## Mumps

## Rubella



# Recent reduction in vaccine uptake impacts on population incidence

- RSC data for 2024 showed a rise in measles (and pertussis) in the last year
- Lower immunisation rates
- Possibly post pandemic factors



## Clinical “coding” in primary care describes the process of care

### Coding is a method of unambiguous recording

- **Avoids “short context issues”:** Diagnosis/key symptoms infer meaning/ communicate
- **Enables hierarchical classification:** Hierarchies based on diseases or symptoms
- **Efficient:** Small volumes of data and processing
- **Data can be analysed at scale:** coded can be readily aggregated and analysed

### History of UK coding:

*Family tree-like – Codes only have one parent – e.g. H, H3, H33*

- **1983: Read codes version 1** (4-byte set) – extended to 30K codes
- **1988: Read codes version 2** (5-byte set) – supported by BMA & RCGP 100K

*Conceptual structures – clinical terms have multiple parents – e.g. disorder & body structure*

- **1994: Read version 3, Clinical Terms (CTv3)** – 200K terms
- **1999: SNOMED-CT** – Systematized Nomenclature for Medicine – Clinical Terms
  - Over 1 million descriptions linked to 360K clinical concepts
  - dm+d (Dictionary of medicines and devices) developed separately but now integrated
  - Captures sociodemographic risk, comorbidities & the process of care + primary care outcomes

# Hospital & death inform the “outcomes of care” data

- **Data linkage enables outcomes of care to be measured**

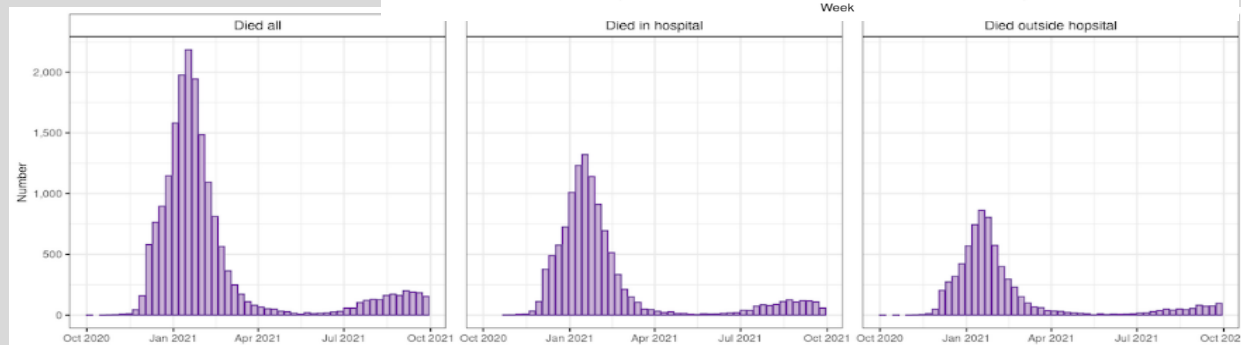
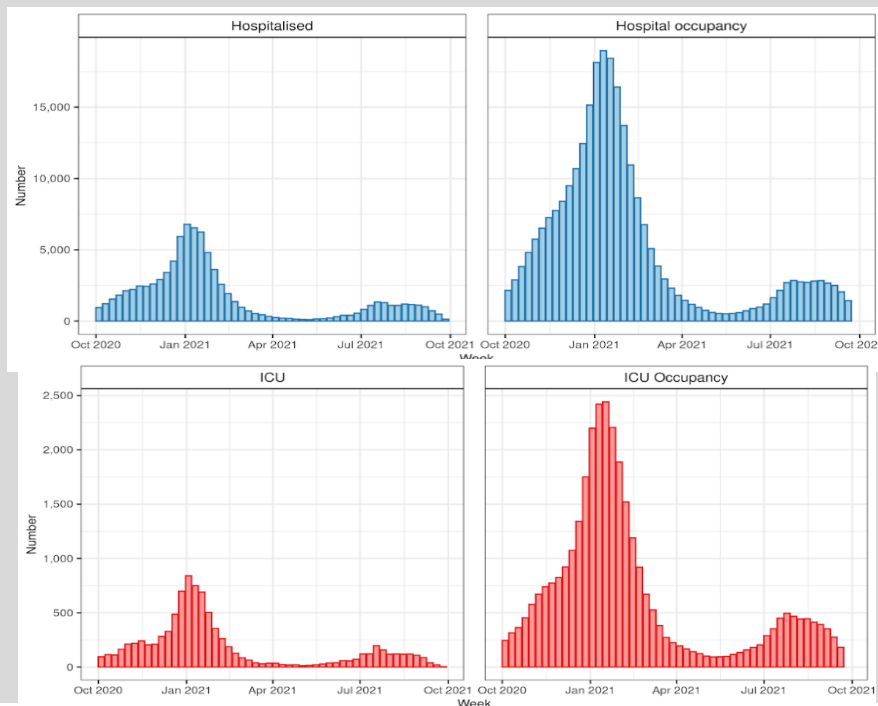
We can link primary care data to hospital admission, intensive care (ICU) & mortality data using pseudonymised NHS number

- **Health outcomes from COVID**

Outcomes of interest include hospitalisation, intensive care, death

**Reference:**

DOI: [10.2196/39141](https://doi.org/10.2196/39141)





# Assessing if our data are of sufficient quality?

## Data quality definitions:

- **Mathematical:** Completeness, accuracy currency
- **Functional:** Fit for purpose, can you infer meaning?

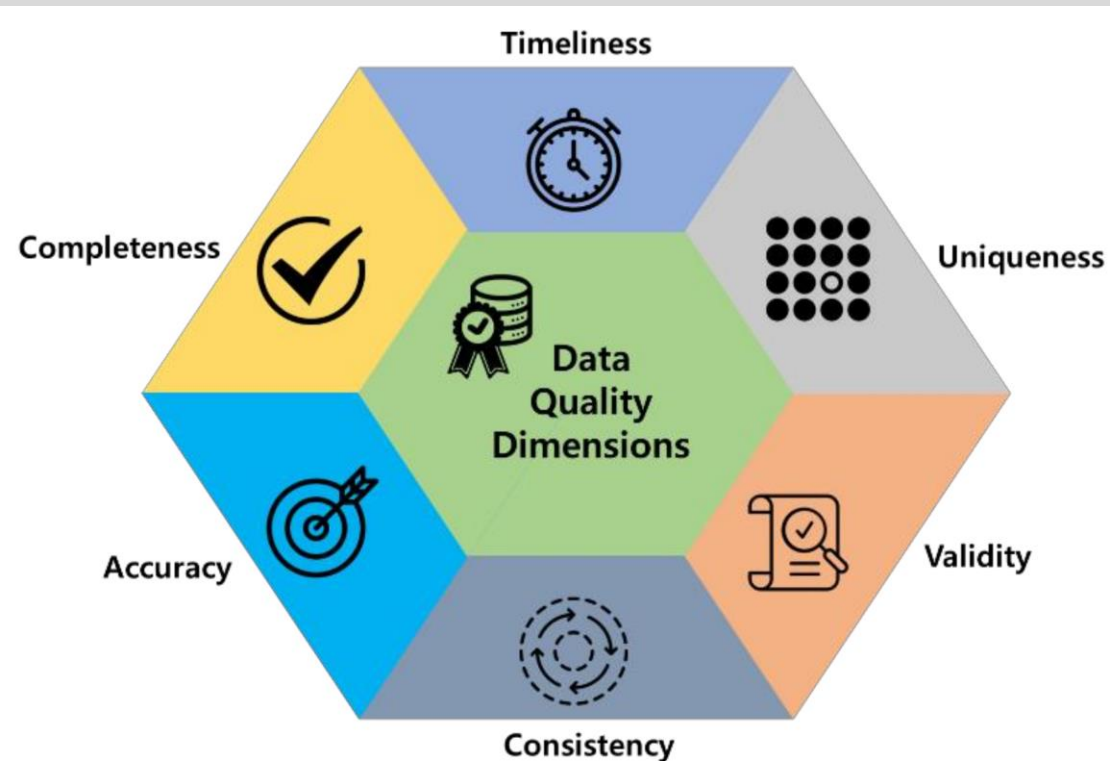
## UK Government now sets out Dimensions of data quality:

- Based on Data Management Association of the UK (DAMA UK)

## References:

DOI: [10.14236/jhi.v14i3.625](https://doi.org/10.14236/jhi.v14i3.625)

<https://www.gov.uk/government/news/meet-the-data-quality-dimensions>



# Strengths & limitations of our data for population health

## **Population:**

Registration-based system + unique identifier

Denominator inflation

Comprehensive sociodemographic (including ethnicity & socioeconomic status data (Equity)), comorbidity, and other measures of risk

## **Interventions & exposures:**

Can be recorded well

Near complete exposure to therapy

## **Comparators:**

Matching can be a challenge

## **Outcomes:**

Many important outcomes are not recorded in primary care data

Much health service management focusses on process rather than outcomes

Managing population health can't be seen in isolation from other health system goals...

## **References:**

DOI: DOI: [10.1093/fampra/cmi057](https://doi.org/10.1093/fampra/cmi057) <https://www.cochranelibrary.com/about-pico>

**P** = Population

**I/E** = Intervention  
/ Exposure

**C** = Comparator

**O** = Outcome

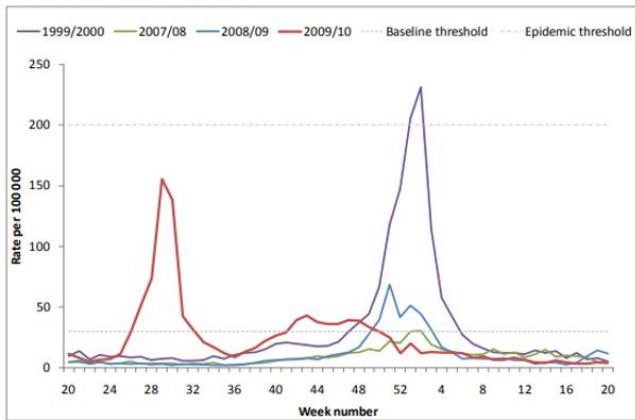
# Summary

1. Population health is a health system objective
2. **Data for surveillance - RSC sentinel network**
3. Summary – Developing the primary care data custodian

# What is sentinel surveillance?

- A “sentinel” watches over or stands guard (a sentry)
  - From old Italian *sentina* - vigilance
  - In our case for anticipated or unanticipated infectious disease
- **Exemplars:** (1) 2009 “Swine-flu” pandemic, (2) COVID-19 community spread (3) First UK Influenza A H1N2v

Figure 8: Royal College of General Practitioners weekly influenza-like illness rate per 100,000, May 2009 – May 2010

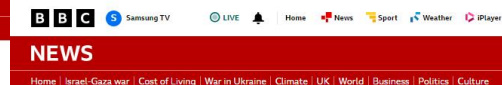


## Coronavirus: Latest patient was first to be infected in UK

© 29 February 2020



Haslemere Health Centre has been closed since Friday morning



## Swine flu: UK detects first human case of new strain

© 4 hours ago



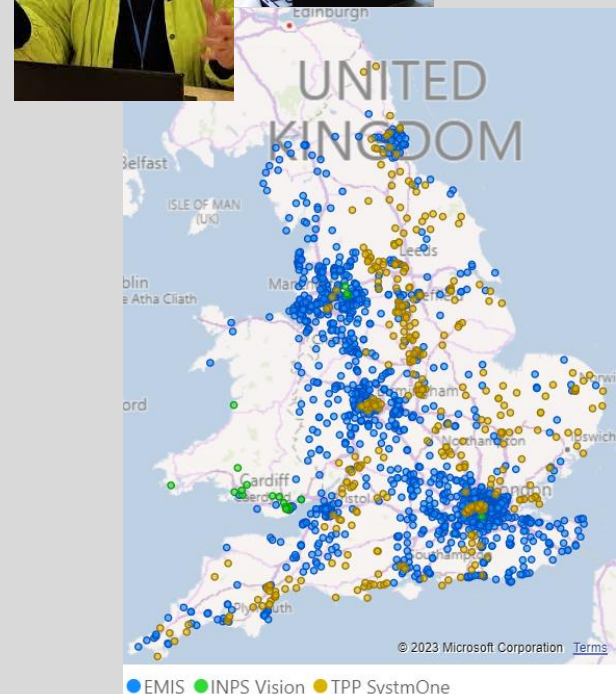
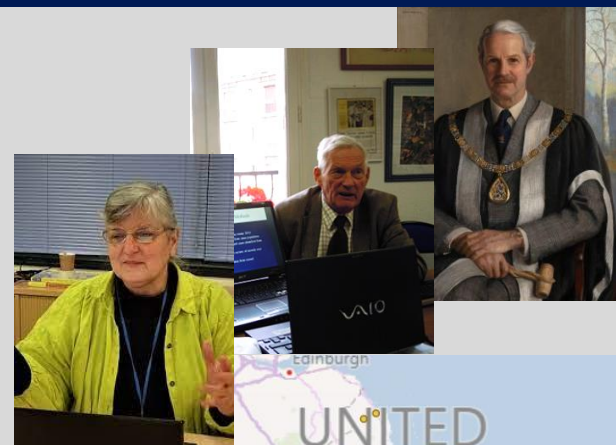
Health officials are investigating the first confirmed case of a new strain of swine flu in the UK.

The A(H1N2)v infection was detected in a routine flu screening test at a GP surgery in North Yorkshire.



# History of the RSC

- One of the **world's oldest sentinel systems**
  - Some data collection – 1957 (RCGP Epidemic Centre)
  - Sentinel surveillance - 1967
  - Virology sampling, seasonal - 1993
  - Serology sampling & all year-round virology – 2020
- **Size** – historically 100 general practices, now >2,000
- **Current size** – **>22 million registered patients**
  - One third of the English national population
  - Data processed in near real time
  - Nationally representative – geography/age/ethnicity/SES
- **Near real-time reporting** – key signals
  - Influenza-like-illness (ILI) – 50% virologically confirmed flu
  - Acute bronchitis/bronchiolitis in <5 year – 40% have RSV

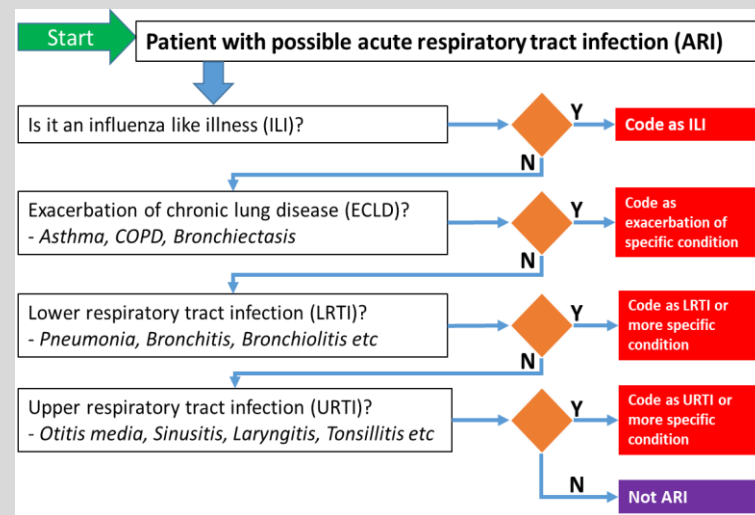




# What does the RSC do?

- Collect pseudonymised data
  - Focus acute respiratory infection (ARI)
    - Is it flu?
    - An exacerbation of chronic lung disease?
    - LRTI or URTI

*Reference:* DOI: [10.2807/1560-7917.ES.2024.29.35.2300682](https://doi.org/10.2807/1560-7917.ES.2024.29.35.2300682)
- Collect **Virology** samples from people with ARI
  - **Goal 1,000 samples per week**
    - 1,080 achieved week 51 2023
  - More virology sampling practices needed
    - Practice & home test options
- Other sampling
  - **Serology** – emphasis on younger patients
  - Asymptomatic testing in children
  - POCT testing
- Reporting
  - Weekly Return for >55years
  - Disease monitoring & vaccine effectiveness
  - Circa 17 million denominator in reports



**Key Statistics:**

Week Number/Year.....	39/2023
Week Starting - Ending.....	25/09/2023 - 01/10/2023
No. of Practices.....	1,526
Population.....	15,036,587

# MONITORY study exemplar of QI and Education

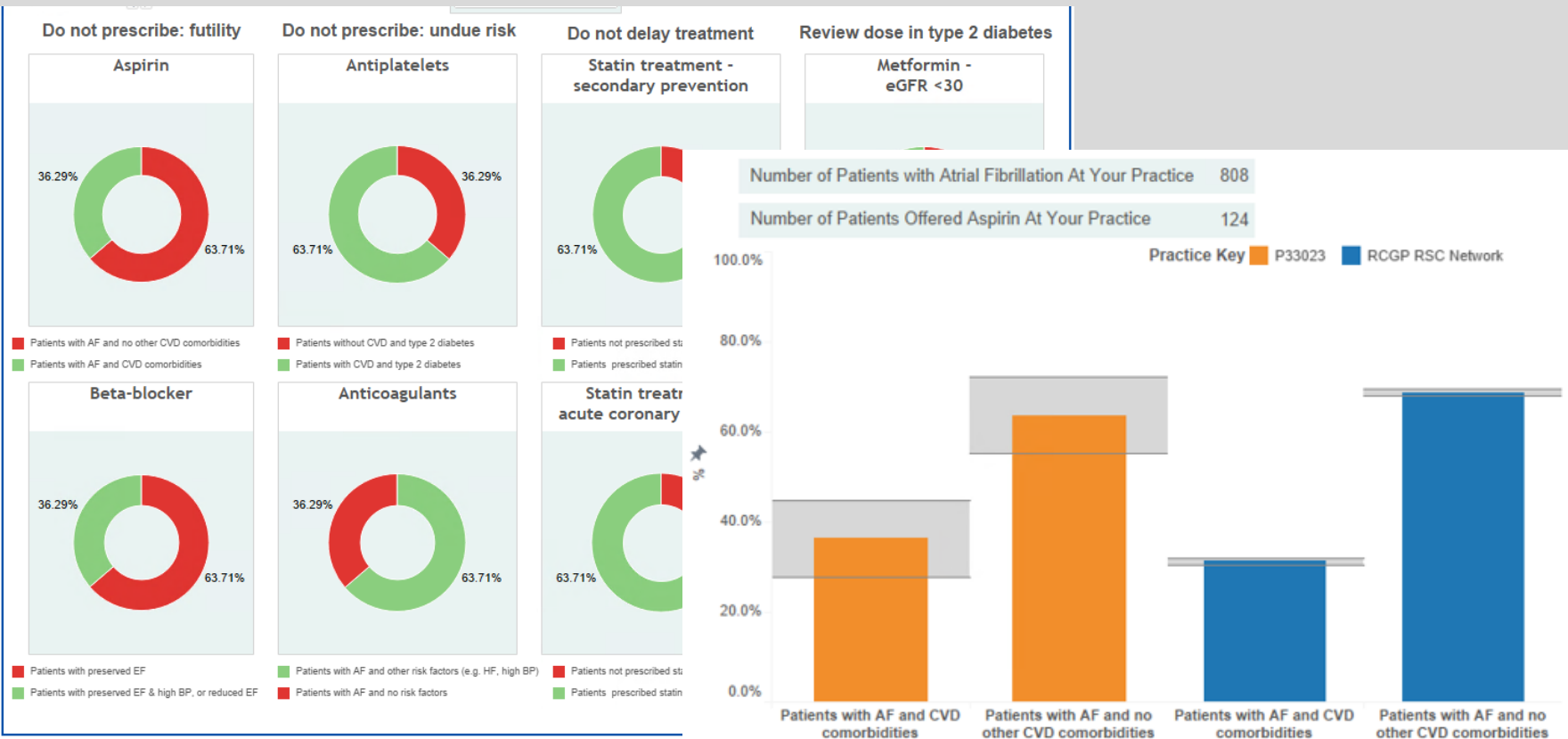
## MONITORY study

- Provides feedback on national guidance & quality of care
- Focus on the management of cardiometabolic disease

**References:**

DOI: [10.2196/52047](https://doi.org/10.2196/52047)

DOI: [10.2196/39141](https://doi.org/10.2196/39141)



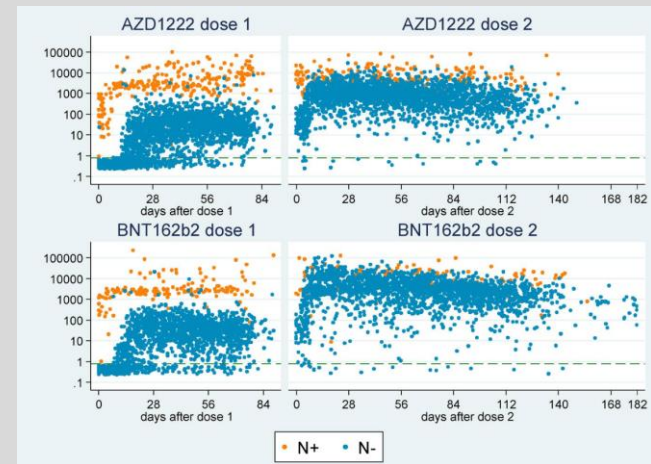
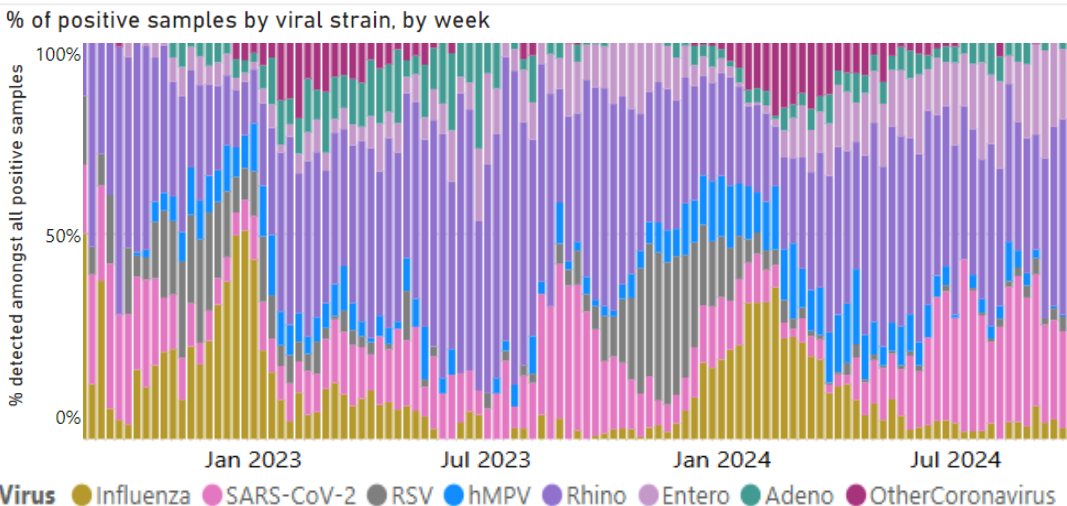
# What's in it for participating practices?

- Contribute to national surveillance / Societal benefit / policy
  - Pandemic contributions – Community spread, Disparities in infection & mortality, Low population immunity (needed a national vaccination programme), Vaccine waning (need boosters)
- Payment for samples taken / requested
- Know aetiology of respiratory disease in your practice
  - Improves antimicrobial stewardship & use of antivirals
- Research study opportunities
- Fantastic practice liaison team to support - [practiceenquiries@phc.ox.ac.uk](mailto:practiceenquiries@phc.ox.ac.uk)

**References:**

DOI: [10.1016/S1473-3099\(23\)00736-3](https://doi.org/10.1016/S1473-3099(23)00736-3)

DOI: [10.1016/S1473-3099\(24\)00352-9](https://doi.org/10.1016/S1473-3099(24)00352-9)



# Summary

1. Population health is a health system objective
2. Data for surveillance - RSC sentinel network
3. Summary – Developing the primary care data custodian

## RSC summary

### RSC

- One of Europe's oldest sentinel networks
- Expanded greatly since 2011, now >22 million patients in contemporary extract
- >2,000 practices sharing data
- Team with a deep understanding of routine data & practice working
- Impact on policy driven by access to data and high-quality analyses – (1) Spread of COVID, (2) Vaccine waning in immunocompromised, (3) European guidance on adherence in T2DM

**References:**

DOI: [10.3399/bjgp22X720965](https://doi.org/10.3399/bjgp22X720965)

DOI: [10.1016/S1473-3099\(23\)00736-3](https://doi.org/10.1016/S1473-3099(23)00736-3)

DOI: [10.1016/S1473-3099\(24\)00352-9](https://doi.org/10.1016/S1473-3099(24)00352-9)

### Future impact on public health and policy

- Work within national policy and context, in the UK migrating the RSC to a secure national data environment and meeting surveillance needs for the future
- Look to ensure we have internationally generalisable solutions
  - International collaborations – *more challenging post BREXIT!*
  - Involvement in European (e.g. EFMI, WONCA Europe) and international (e.g. IMIA, WONCA) organisation
- International agreement around validation of methods and data



# Population health can't work in isolation from health system goals

- Population health
- Patient experience
- Reducing (controlling) costs
- Care team wellbeing
- Equity & Inclusion



*References:*

DOI: [10.1001/jama.2021.25181](https://doi.org/10.1001/jama.2021.25181)

<https://doi.org/10.1007/s11606-021-06846-x>

## England has world leading national health data systems

- Unique ID - NHS number
  - Gives a population denominator /know the rate of disease
  - Links data across healthcare (including Apps)
  - Can monitor inequities in health care
- Comprehensive national records
  - GP CMR systems – used in-consultation
  - Emergency services are largely provided by the NHS
  - Some elective care in the private sector
- National data collections
  - Use of SNOMED clinical terms - machine processable
- Willingness of GPs to share (pseudonymised) data
  - For quality improvement & research
  - Over 2,000 practices kindly share data with the RCGP Research & Surveillance Centre (RSC) – working primarily to support public health with UKHSA
- Should GPs remain data controllers?
  - Sudlow review says UK-wide data sharing is not possible
  - Given the right contractual framework 100% GP uptake – e.g. QOF, PCNs (in England)



*References:*

DOI: [10.1093/fampra/ctmi106](https://doi.org/10.1093/fampra/ctmi106)

DOI: [10.2196/39141](https://doi.org/10.2196/39141)

DOI: [10.3399/bjgp22X720965](https://doi.org/10.3399/bjgp22X720965)

# Misalignment of information system with goals of the health system

## *Misalignment (with our health system aims):*

- Patient care tools that don't measure access or workload
  - Primary care workload & pressures are largely unrecorded
- Much clinical work in primary care is not monitored
  - Results (blood and other tests)
  - Reviewing hospital letters and imaging results
  - Prescriptions
  - Triage
  - Other clinical tasks
- Current focus is on the process of care
  - Lack of focus on health outcomes (population health)
  - Access may not be key to experience or for outcomes...



## *Changes needed:*

- Strategic focus on interface between technologies & organisations
  - Patient information portals that are independent of the CMR
  - NHS 111 messages that are not coded
- Training to incorporate technology into clinical workflow
  - Coding is caring – Data quality
  - Learning outcome for all going into primary care
  - “Data quality” / Data custodian training should be part of primary care
- Research
  - How to use technology & reduce staff time
  - World leading primary care research - but a stressed workforce!

## Access to health records will continue to build

- Generally welcomed
  - Misalignment of record design & new purpose
  - From clinician aide memoire to patient information
  - No issues with sharing in overwhelming majority
  - May help improve patient experience & access to information
  - Scope to improve the evidence base
- NHS App and others have provided a usual tool
  - Results
  - Prescriptions
  - COVID-19 vaccinations & certificates
  - Symptom checker
- In some risk of harm
  - Certain vulnerable groups
  - Possible impact on attendance

*References:*

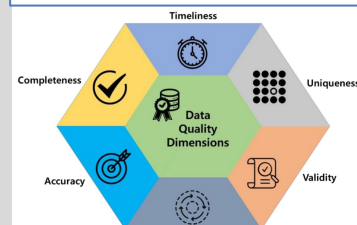
DOI: [10.2196/13042](https://doi.org/10.2196/13042)

•DOI: [10.1055/s-0038-1641202](https://doi.org/10.1055/s-0038-1641202)



# Data quality is the fundamental building block of measuring quality

- Data quality is fundamental building block
  - Great systems for patient care/population health, and equity
  - Much weaker for analysing patient experience, cost workforce and workload issues
- What we already have is internationally renowned – however time to step up!
- Bottom-up professional-led is best
  - Preferred solution for having the data quality is a contractual change that funds “Data Custodianship”
  - The Sudlow review provides an opportunity to discuss further
- What might primary care “Data Custodianship” look like?
  - Needs to be part of our contract and funded
  - Data quality learning outcomes in training for all working in primary care
  - Managing risks associated with GDPR (not managed by the NHS)
  - Elements of data curation
    - (1) Historic record curation – preparing for access
    - (2) Appropriate access (for patients, carers + other HCPs)
    - (3) Data quality - “Coding is caring” + flagging risk + metadata
    - (4) Transfer / sharing records for care & where consent for research
- Coding is caring!





**Clinical coding & wider data quality are essential for  
population health**

**Oxford-RCGP Research & Surveillance Centre (RSC)**

***Thanks for listening***

**Simon de Lusignan**  
Nuffield Dept Primary Care Health Sciences

**@Lusignan\_S**

**[simon.delusignan@phc.ox.ac.uk](mailto:simon.delusignan@phc.ox.ac.uk)**

