Opportunities & challenges of using routine clinical data – *a.k.a. Real World Evidence* for surveillance, quality improvement & research

Simon de Lusignan

Professor of Primary Care & Clinical Informatics

Head of Department of Clinical & Experimental Medicine

s.lusignan@surrey.ac.uk

www.surrey.ac.uk

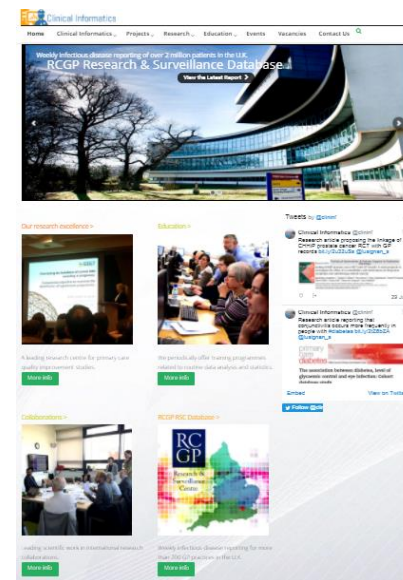# My practice & academic work
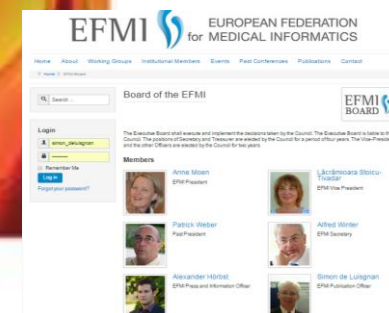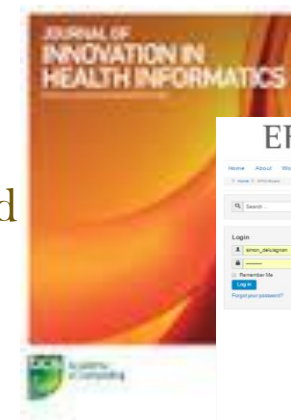
- ## GP in Guildford >30years
  - 12,000 patient practice - 10 GPs - 4 partners
  - Computerised since 1988 – EMIS brand since 1994

- ## Professor of Primary Care & Clinical Informatics

- ## Head of Dept Clinical & Experimental Medicine
  - Clinical Informatics & Health Outcomes Research Group: *How IT can be used to improve quality*
  - Caldicott Guardian for University

- (1) Routine data for surveillance
- (2) Using Real World Evidence (RWE)
- (3) Scholarship:

  - Chair PCI WG, UK Rep & Publications officer EFMI Board member & Chair IMIA PCI WG
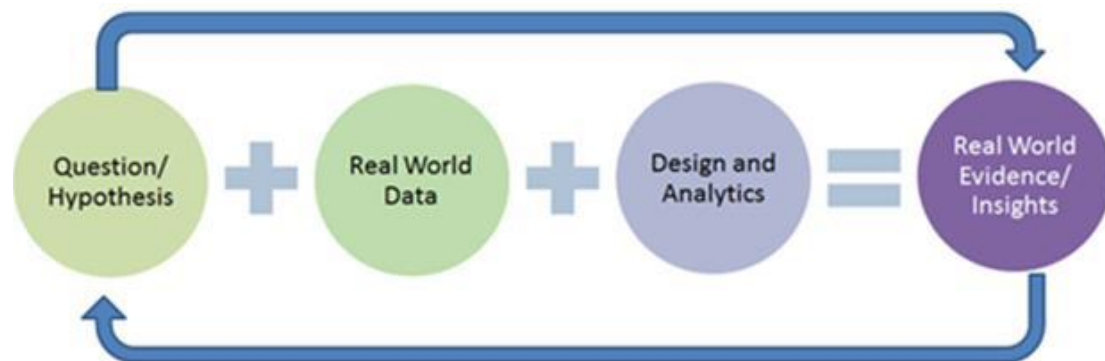 - Editor of Journal of Innovation in Health Informatics

Real World Evidence (RWE) for SQIR

- **Real world evidence (RWE)**
- Nationally available data for RWE studies
- RCGP RSC
  - History & Role
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine Effectiveness
- Quality Improvement & Research
- Conclusion:
  - Scope for collaborative research

Scope of my **RWE** research:

- **CMR** contain useful data for:
  - Epidemiology
  - Medication use & Effectiveness
  - Trials in routine practice
- **Other big data**
  - Social Media (Twitter)
  - IoT (Older people)
  - Device data
- **We need to better define real world evidence (RWE) studies**

Primary Care Computerised Medical Record (CMR)

Pathology
- Microbiology
- Haematology
- Chemical Pathology

Phenotype/ Morbidity / Disease
- Primary care record

Medicine & Vaccine exposure
Nearly complete

Question/ Hypothesis + Real World Data + Design and Analytics = Real World Evidence/ Insights

Routine data – a.k.a. RWE

- Opportunities:
  - Lots of it
  - Value increased when linked to other data

- Challenges – interpretation & methodology
  1. Consistency of key data definitions
  2. Validation & extension of RWE studies
  3. Properly theorised approach to QI
  4. Segmenting the population & applying risk scores

Real World Evidence (RWE) for SQIR

- Real world evidence (RWE)
- Nationally available data for RWE studies (UK)
- RCGP RSC
  - History & Role
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine Effectiveness
- Quality Improvement & Research
- Conclusion:
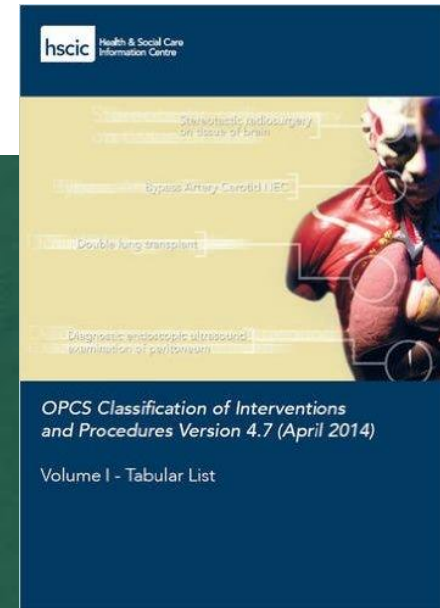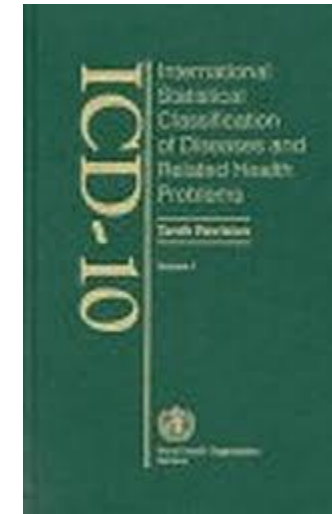 - Scope for collaborative research

National & service data sources

- ## National data
  - NHS Digital https://digital.nhs.uk/
  National Statistics www.gov.uk
  - Neighbourhood statistics
  www.neighbourhood.statistics.gov.uk/

- ## General Practice / Primary Care
  - Single brand of GP computer systems
  - *EMIS – QResearch; Vision (& EMIS) – CPRD & THIN; TPP – ResearchOne*

  - All brands of GP CMR
  - *Royal College of General Practitioners (RCGP) Research & Surveillance Centre (RSC)*

- ## Secondary care /Hospital data:
  - Acute hospitals – Hospital Episode Statistics (HES)
  http://content.digital.nhs.uk/hes
  Mental Health
  http://content.digital.nhs.uk/mentalhealth
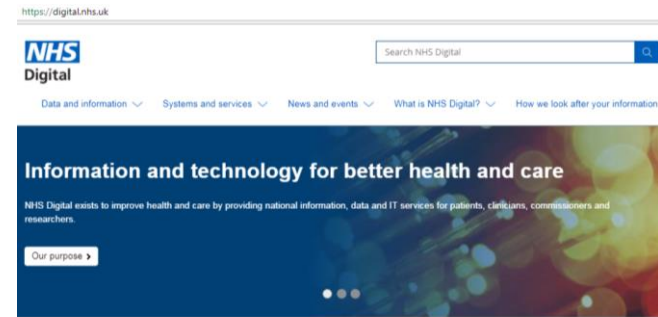
How data are recorded

- ## Hospital data (Episodes)
  - ICD-10 for diseases
  - OPCS-4 for procedures

- ## Primary care data (Encounters)
  - Read codes
    - Version 2 hierarchical
    - CTv3

- ## Move to implement a comprehensive coding system
  - SNOMED CT
  - DM+D UK Drug Extension or Dictionary of Medicines https://www.nhsbsa.nhs.uk/

hscic Health & Social Care Information Centre

OPCS Classification of Interventions and Procedures Version 4.7 (April 2014)

Volume I - Tabular List

ICD-10

International Statistical Classification of Diseases and Related Health Problems

Tenth Revision

SNOMED CT
The global language of healthcare

Real World Evidence (RWE) for SQIR

- **Real world evidence (RWE)**
- **Nationally available data for RWE studies**
- **RCGP RSC**
  - **History & Role**
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine Effectiveness
- **Quality Improvement & Research**
- **Conclusion:**
  - Scope for collaborative research

RCGP RSC @50 years

- One of the longest established primary care sentinel network
- Since 2015, new data & analytics hub at University of Surrey
- Data is extracted on a weekly basis from a network of

  >170 GP practices comprising >2 million patients.

- Principal output:

  - Weekly surveillance report of infectious and respiratory diseases -
  - Used by Public Health England
  - Data informs the Chief Medical Officer when flu is circulating

- The practices in the network receive continuous feedback – Audit Based education
- Online at: www.rcgp.org.uk/rsc
- Sign up to receive our "Weekly Return" produced for >50-yeaers

RCGP RSC History

- 1957 first research centre:
  - Records and Statistical Unit
- 1967 Birmingham Research Unit (BRU)
  - "Weekly returns" for >50 years
  - Founder Donald Crombie
- 1989 Douglas Fleming - Director
  - Gold standard sentinel network

- 2013 – appointed Director
  - Data & analytics hub @ University of Surrey since March 2015

BIRMINGHAM RESEARCH UNIT

Weekly report on surveillance of communicable and respiratory diseases

– www.rcgp.org.uk/rsc

Royal College of General Practitioners
Research and Surveillance Centre

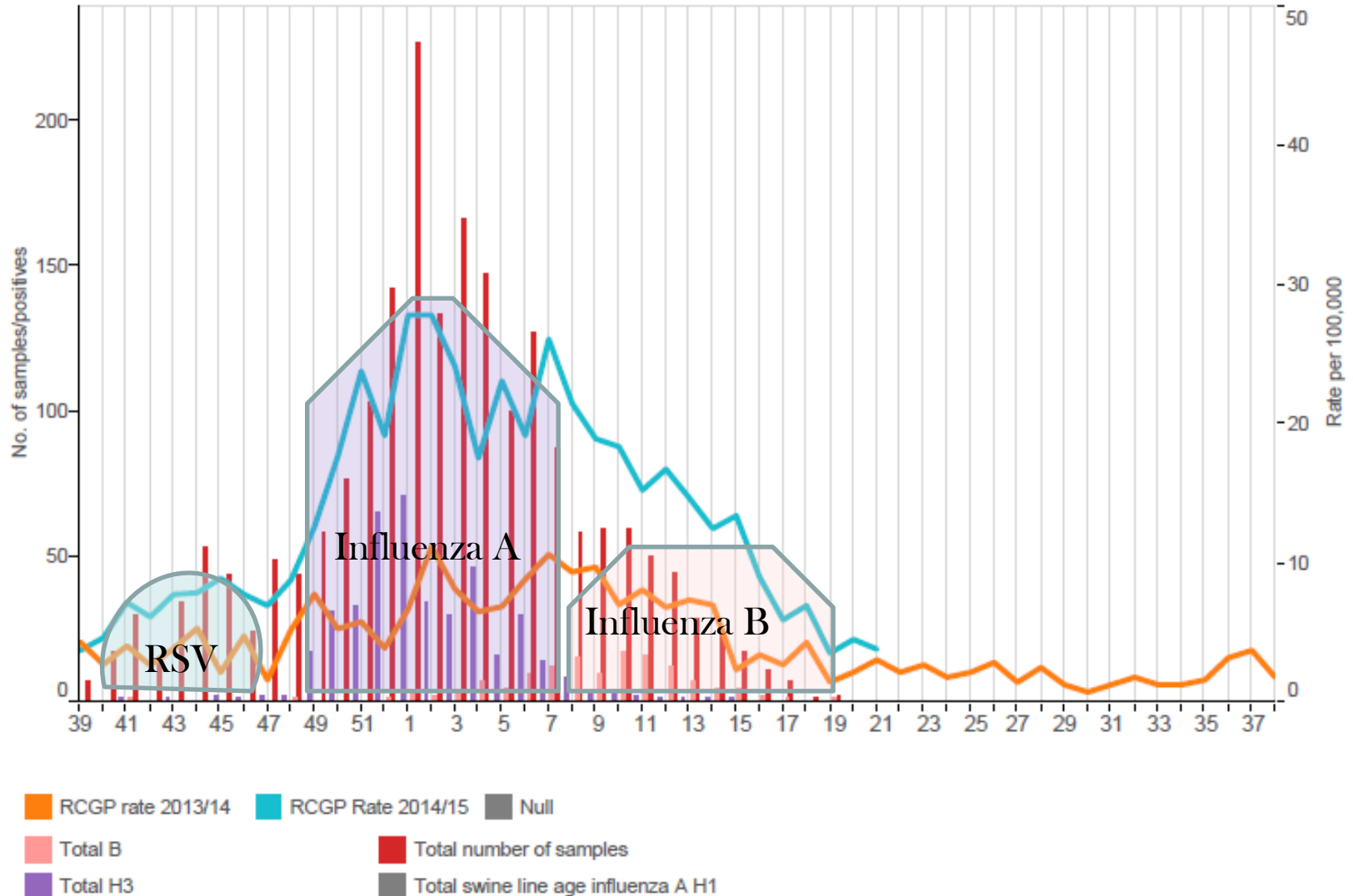## RSC Communicable and Respiratory Disease Report for England

Key Statistics:

Week Number/Year................19/2017
Week Starting - Ending...........08/05/2017 - 14/05/2017
No. of Practices.....................153
Population..............................1521333

National (England)

- **Acute Bronchitis** : increased from **47.6** in week 18 to **57.0** in week 19.
- **Asthma** : increased from **11.6** in week 18 to **12.4** in week 19.
- **Common Cold** : increased from **53.7** in week 18 to **69.2** in week 19.
- **Influenza-Like illness** : increased from **2.6** in week 18 to **3.0** in week 19.
- **Respiratory System Diseases** : increased from **214.7** in week 18 to **256.3** in week 19.

.clininf.eu
surrey.ac.uk

Surveillance of influenza: 2014/2015

RCGP — Royal College of General Practitioners

Influenza A

Influenza B

RSV

No. of samples/positives

Rate per 100,000

Legend:
- RCGP rate 2013/14
- RCGP Rate 2014/15
- Null
- Total B
- Total number of samples
- Total H3
- Total swine line age influenza A H1

www.cnmmn.eu
www.surrey.ac.uk

Surveillance of influenza: 2015/2016

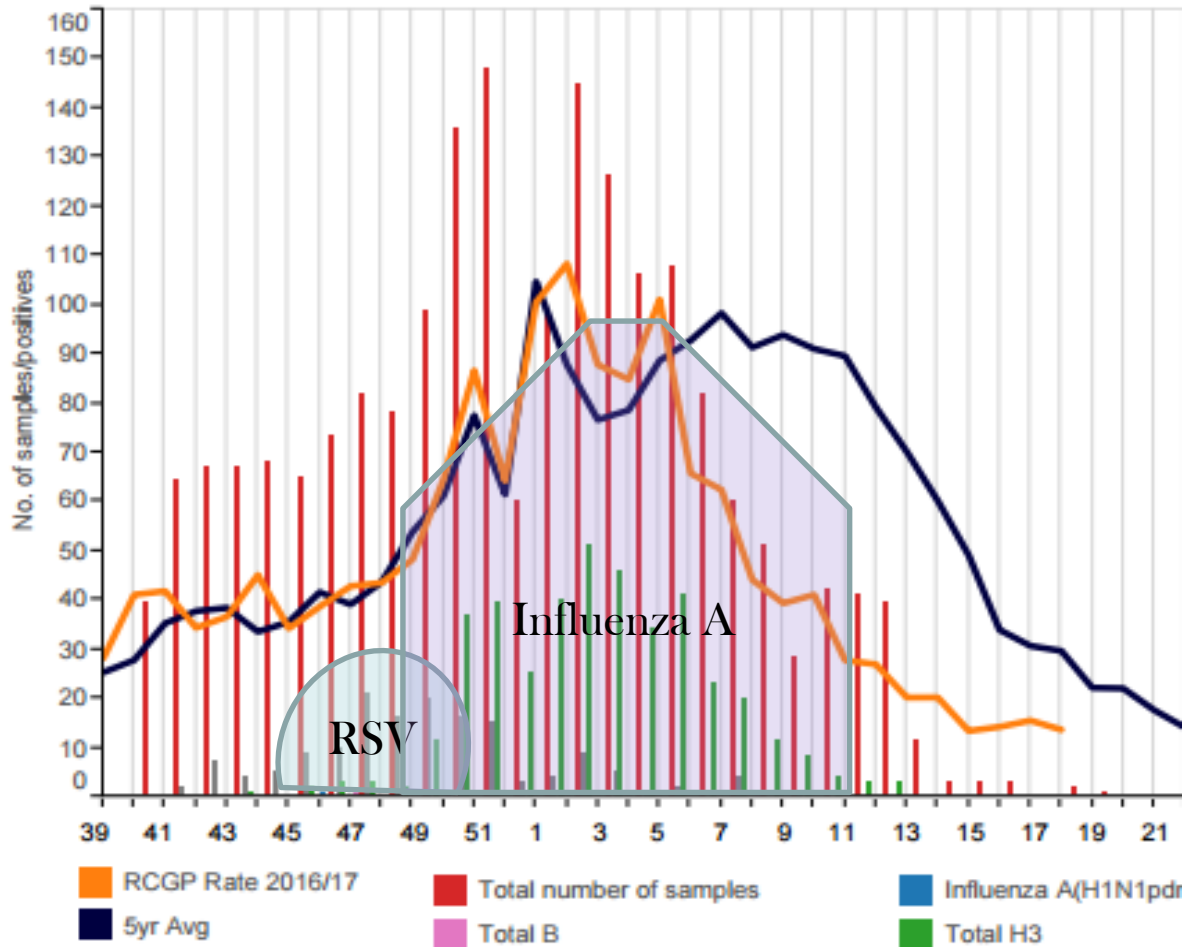(B) RCGP/PHE Influenza Swabbing Surveillance 2015/16 (all ages, gender,

Legend:
- RCGP Rate 2015/16
- 5yr Avg
- Total number of samples
- Total B
- Influenza A(H1N1pdm09)
- Total H3

Labels on chart: RSV, Influenza A, Influenza B
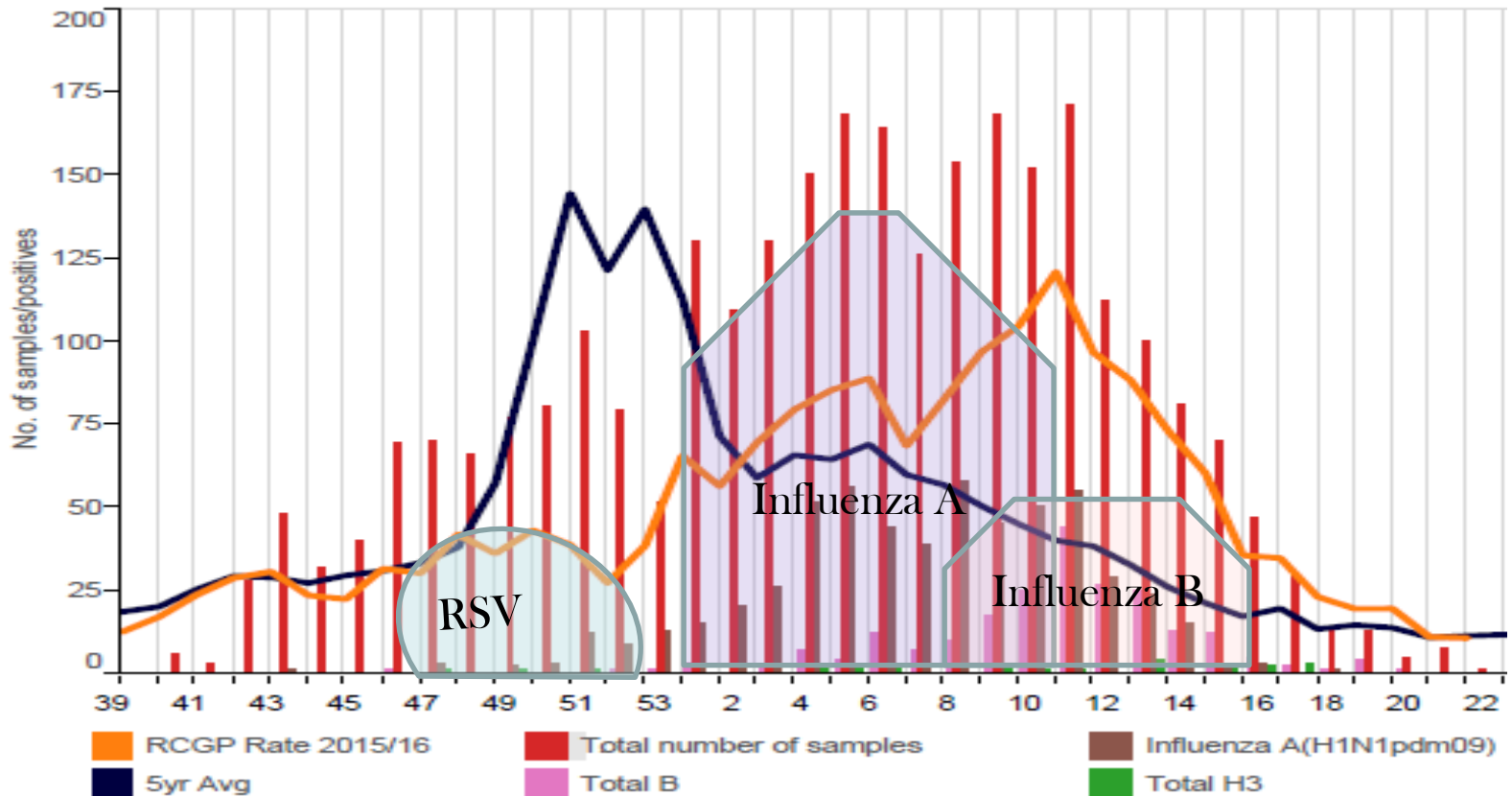
RCGP RSC
– 50ᵗʰ influenza season
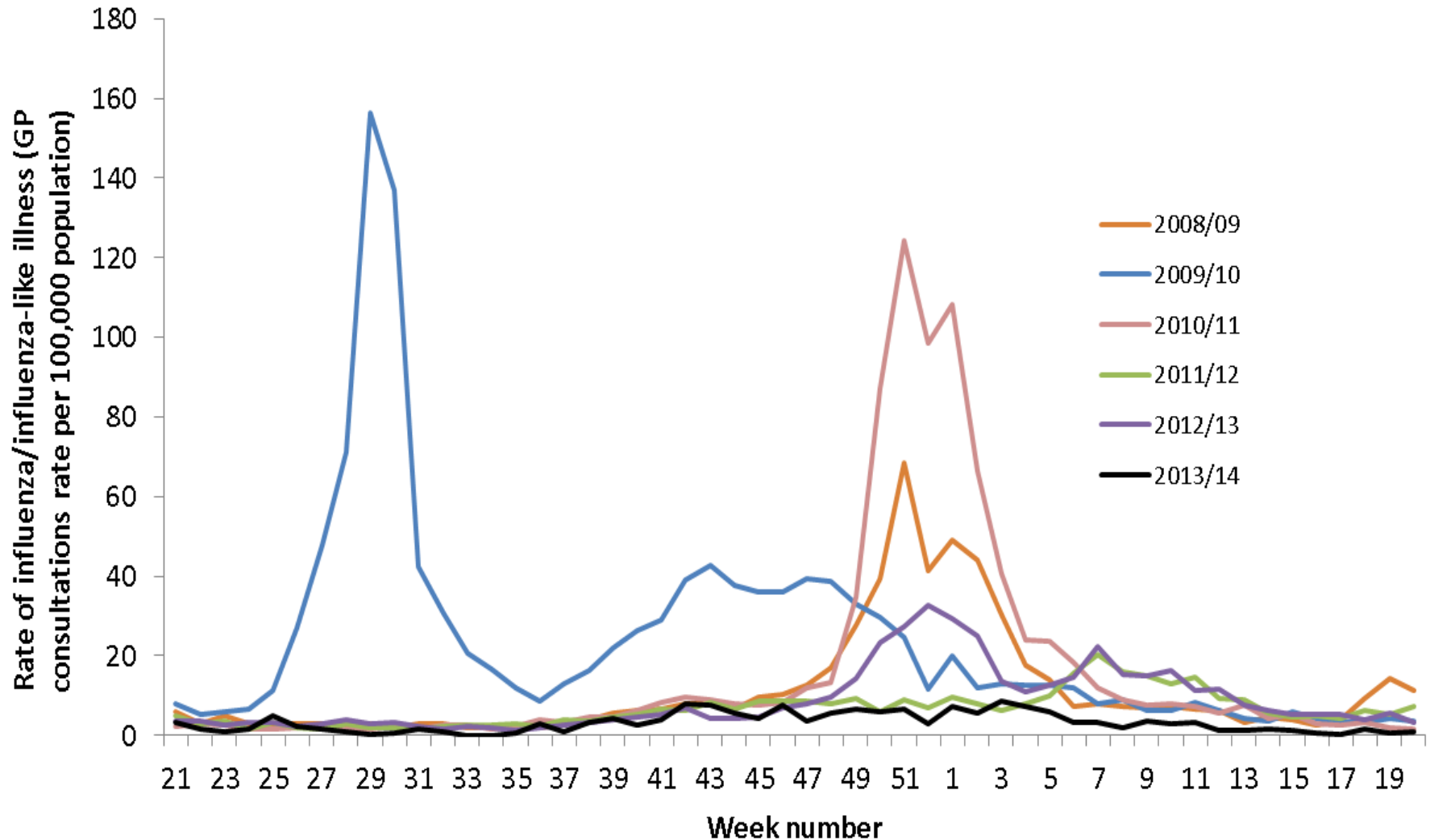
UNIVERSITY OF
SURREY

(B) RCGP/PHE RSV and Influenza Virology Swab Surveillance 2016/17

- **Little influenza B this season!**

RCGP RSC spotted swine flu pandemic 2009/2010

Real World Evidence (RWE) for SQIR

- Real world evidence (RWE)
- Nationally available data for RWE studies
- RCGP RSC
  - History & Role
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine Effectiveness
- Quality Improvement & Research
- Conclusion:
  - There is greater capacity for collaborative research

Importance of data linkage

- **Donabedian's evaluation framework**
  - Structure
  - Process
  - Outcome

- **In the UK:**
  - Structural data about workforce & facilities if often challenging to obtain
  - Primary care data provides information about process
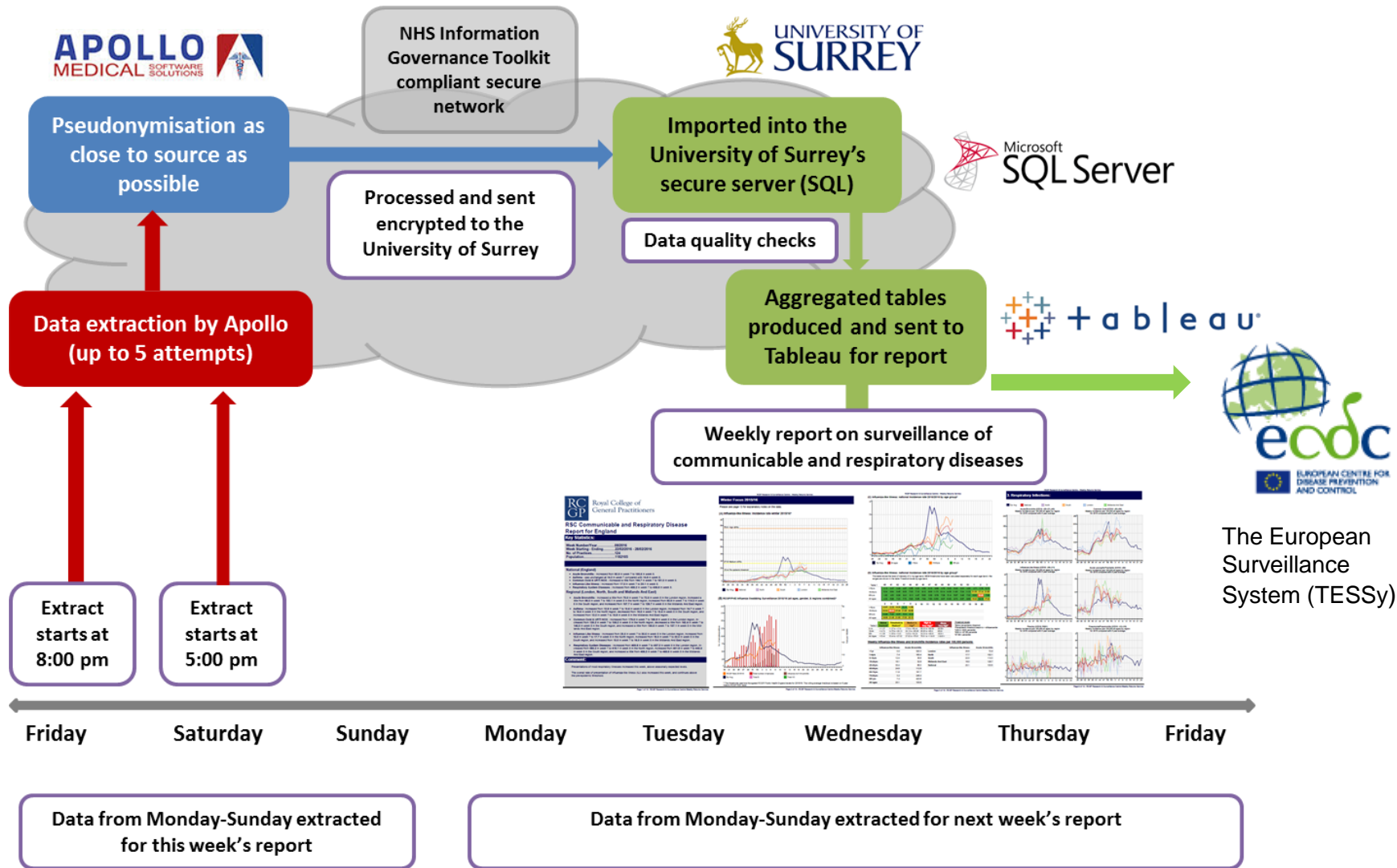 - Hospital / ONS data about outcomes

Capability for data linkage:

- The **RCGP RSC** is capably of "pseudonymised" linkage

- Links to other data – on project-by-project basis
  - Hospital Episode Statistics – HES
  - Cancer Registry
  - Mental Health data

- Capability within the group to link to other data sources
  - MINAP
  - SINAP
  - IAPT (where no NHS number))
  - Renal Registry

Weekly data extraction process

Pseudonymisation as close to source as possible

NHS Information Governance Toolkit compliant secure network

Data extraction by Apollo (up to 5 attempts)

Processed and sent encrypted to the University of Surrey

Imported into the University of Surrey's secure server (SQL)

Data quality checks

Aggregated tables produced and sent to Tableau for report

Weekly report on surveillance of communicable and respiratory diseases

Extract starts at 8:00 pm

Extract starts at 5:00 pm

Friday   Saturday   Sunday   Monday   Tuesday   Wednesday   Thursday   Friday

Data from Monday-Sunday extracted for this week's report

Data from Monday-Sunday extracted for next week's report

The European Surveillance System (TESSy)
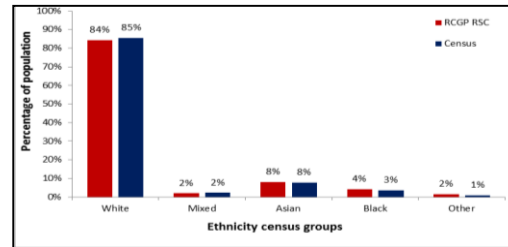
www.surrey.ac.uk

Real World Evidence (RWE) for SQIR

- Real world evidence (RWE)
- Nationally available data for RWE studies
- RCGP RSC
  - History & Role
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine Effectiveness
- Quality Improvement & Research
- Conclusion:
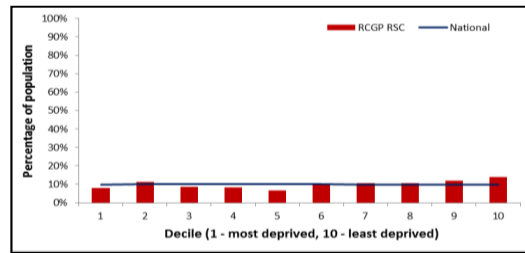 - There is greater capacity for collaborative research
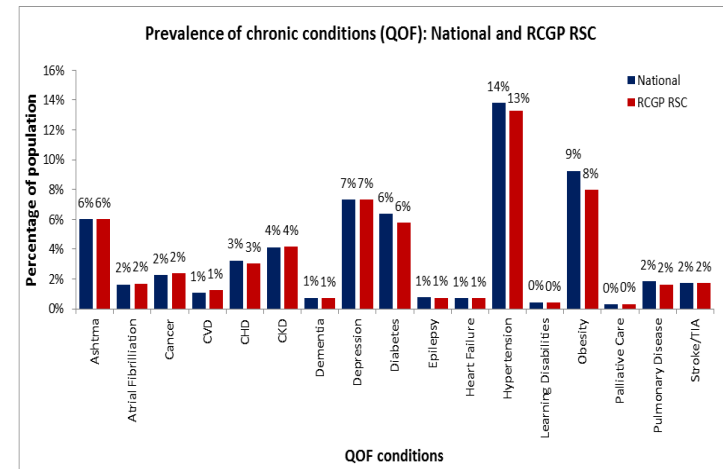
Representativeness of RCGP RSC
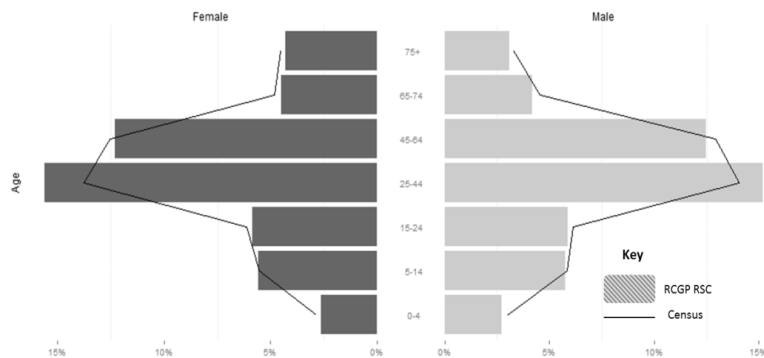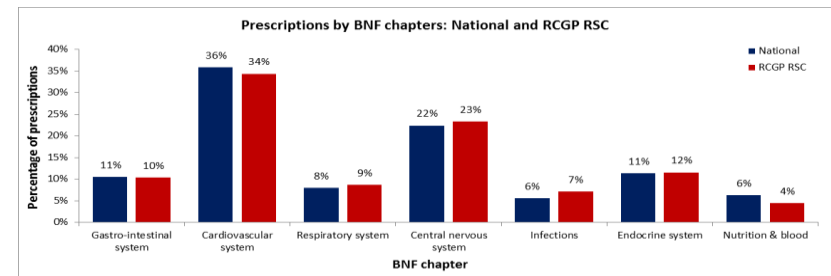
Geographical distribution

Ethnicity

Deprivation

Chronic disease

Age-sex profile

Prescribing

www.surrey.ac.uk

Summary of representativeness

- The RCGP RSC is representative of the English population on a variety of domains, both demographic and clinical.

- Fresh data – uploaded twice per week

- Practices who will take microbiological samples & conduct questionnaires

Real World Evidence (RWE) for SQIR

- Real world evidence (RWE)
- Nationally available data for RWE studies
- RCGP RSC
  - History & Role
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine effectiveness research
- Quality Improvement & Research
- Conclusion:
  - There is greater capacity for collaborative research

Flu vaccine effectiveness:

- **Why is it important**
  1. Excess winter mortality – flu vaccine coverage & mismatch major factor

  2. Effectiveness of vaccinating children

- **Test negative case control (TNCC)**
  - Standard approach
  - Compares - Vaccine exposed (Y/N) with Confirmed case (Y/N)

www.clininf.eu
www.surrey.ac.uk

Real World Evidence (RWE) for SQIR

- Real world evidence (RWE)
- Nationally available data for RWE studies
- RCGP RSC
  - History & Role
  - Structure & Capability for linkage
  - Dataset & representativeness
  - Vaccine Effectiveness
- Quality Improvement & Research
- Conclusion:
  - There is greater capacity for collaborative research

Quality Improvement & Research

- Overview of RWE highlights & Current research (since 2015)

- Challenges:
  1. Consistency of key data definitions
     - Ontological approach
     - Health system change / contextual impact on data recording
  2. Validation & extension of RWE studies
     - Validated questionnaires
     - Virology samples
     - Serology & Genetic data
  3. Properly theorised approach to QI
     - Education – Audit-Based Education (ABE)
     - Behavioural change model – COM-B model
  4. Segmenting the population & applying risk scores
     1. Few segmentation tools – e.g. John Hopkins & Ambulatory care groups
     2. Mover from generic risk to modifiable
     – e.g. Risk of admission vs. Risk of Falling

www.clininf.eu
www.surrey.ac.uk

Overview of current RWE research:

- ## Programme level work

  Royal College of General Practitioners (RCGP) Research & Surveillance Centre (RSC), England's primary care surveillance system

  Eli-Lilly, European Real World Evidence (RWE) centre for Type 2 Diabetes – adherence, treatment thresholds (not brand specific), disparities, progression

  Living lab – Surrey & Boarders NHS Trust led study

- ## European awards

  Vaccine research: AVANCE ADVANCE   FluCop FLUCOP   i-Move, YETI

  Others: MOCHA (child health)   iHealth-D2 (diabetes) HealthPros

- ## UK awards

  Integrate (DH-Welcome challenge fund) – gastroenteritis

  London Life Sciences – heart disease risk in British Asians

  Renal anaemia in primary care

  Underlying rate of renal function in routine data

  European Medicines Agency side-effects monitoring in real time

  Amputation study – predicting risk of lower limb disease in DM

  Surrey Health Partners: Primary Care CAG & Surrey cohort

RWE Research highlights (2015 onwards)

Clinical Informatics/Diabetes:

Best research observing health outcomes in routine data

- Nature. 2017;541(7635):81-86;
- Diabetes Care. 2017 Mar;40(3):e30-e31. doi: 10.2337/dc16-2320
- J Med Internet Res. 2016 Nov 25;18(11):e310
- Lancet Diabetes Endocrinol. 2016;4(11):888-889
- Lancet Diabetes Endocrinol. 2015;3(7):526-34.

Flu vaccine effectiveness:

In season research about epidemics vaccine effectiveness, extending into serology

- Euro Surveill. 2016;21(38)
- Euro Surveill. 2016;21(13)
- Lancet. 2017 Jan 21;389(10066):250-251. doi: 10.1016/S0140-6736(17)30046-6

Primary Care Research

Wide range of project based studies - expertise in Physician Associate research:

- PLoS Med. 2017 Mar 14;14(3):e1002252. doi: 10.1371/journal.pmed.1002252
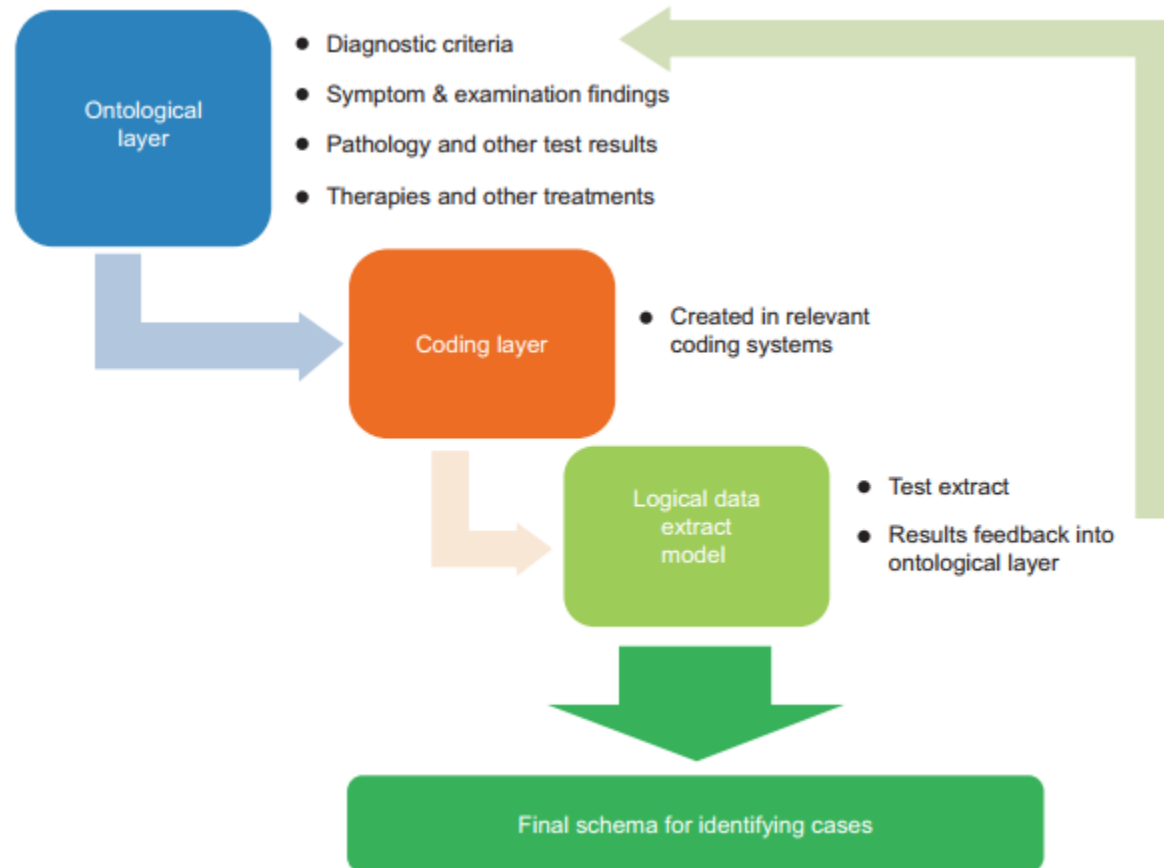- Br J Gen Pract. 2017 Jan;67(654):e29-e40. doi: 10.3399/bjgp16X688573

Challenges: 1. *Consistent definitions*
Ontological approach

- Problem – with opaque case definitions and key outcomes
  - Codes selected to write papers was often opaque
  - Hard to reproduce internationally

- "Solution"
– an ontological approach
  - Transparent mapping from clinical concept to coding list
  - Plus include and document testing
  - Formal recording of ontology in OWL

Ontological layer
- Diagnostic criteria
- Symptom & examination findings
- Pathology and other test results
- Therapies and other treatments

Coding layer
- Created in relevant coding systems

Logical data extract model
- Test extract
- Results feedback into ontological layer

Final schema for identifying cases

Challenge 1: SNOMED CT
Coding system change in England

- Last release of Read was in April 2016

- Primary care in England to transition to Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) April 2018
  - Problems of searching a poly-hierarchical system

- The Dictionary of Medicines and Devices (dm+d) system of coding drugs has been using SNOMED CT terminology

Challenge 1: Organisational change
Workforce & Organisational

- ## Much more mobile primary care workforce
  - Over-regulation & litigation rules makes GP unattractive
  - Creaking health & social care system
  - Older GPs are leaving
  - More salaried doctors
  - Locums are better paid

- ## Practices are being encouraged to amalgamate

UK doctors in Australia – Why they won't be going home any time soon

© 05/03/2015   👤 Edwin Kruys

Doctors who left NHS for Australia: 'More cash, fewer hours, less pressure'

As the row over junior doctors' contracts intensifies, thousands of NHS staff are applying to work abroad. Here three who have moved explain why they went

ⓘ British junior doctor Joe Marwood outside St Vincent's Hospital in Sydney, Australia. Photograph: Matthew Abbott for the Observer

S arah Wollaston, chair of the Commons health committee, revealed last month that her daughter, a junior doctor, and eight of her friends had all quit the NHS to find work in Australia. It's a disturbing trend; but as the dissatisfaction over hours and conditions in the NHS grows, it may be one we have to get used to.

Australia has long attracted doctors from Britain. Depending on their seniority, doctors can earn up to 50% more in Sydney or Melbourne, despite generally working less overtime. With the health secretary, Jeremy Hunt, now threatening

UK doctors are leaving their country en masse. Image: Pixabay.com

The numbers are telling: About 1,500 UK doctors move to Australia and New Zealand each year. This exodus is causing havoc in England. A GP-shortage creates high workloads and overstretched doctors, and a survey showed that over half of UK GPs plan to retire before the age of sixty. This stressful situation has prompted a coming home campaign to entice doctors to

UNIVERSITY OF
SURREY

- Constant need to be more certain about case definition, collect more data, & extend scope of RWE studies
    - More detailed data about specific medicine (brand of vaccine)
    - Validated questionnaires
    - Virology samples
    - Serology
    - Genetic data

- Exemplars
    - Brand specific vaccine data
    - Serological studies as an exemplar

Challenge 2: better data
Comparing vaccine brands
Real time benefit / risk

Vaccines by brand known/unknown

10.6%

49.8%

39.6%
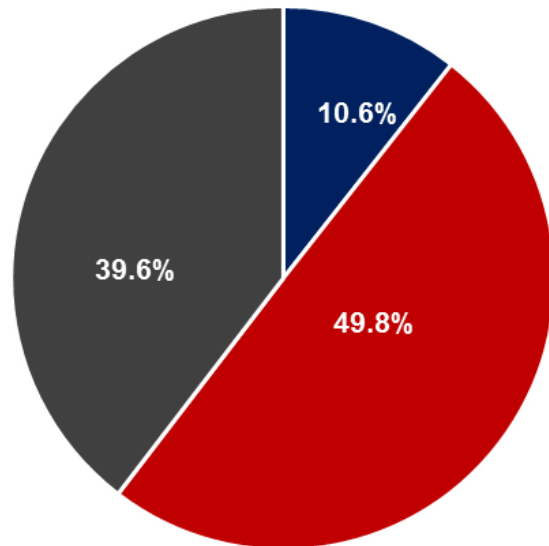
- Brand known (Drug Name)
- Brand known (Batch Number)
- Brand unknown

- Identify the brand of 60.4% of influenza vaccinations
- Using the batch number was better for brand identification
- Specific enquiry during coming seasons may identify a larger %

Brand specific adverse events of interest post vaccination

- Some brands had a significantly larger share of AEIs post vaccination, compared to their vaccine share

Currently unpublished – not for sharing please)

UNIVERSITY OF SURREY
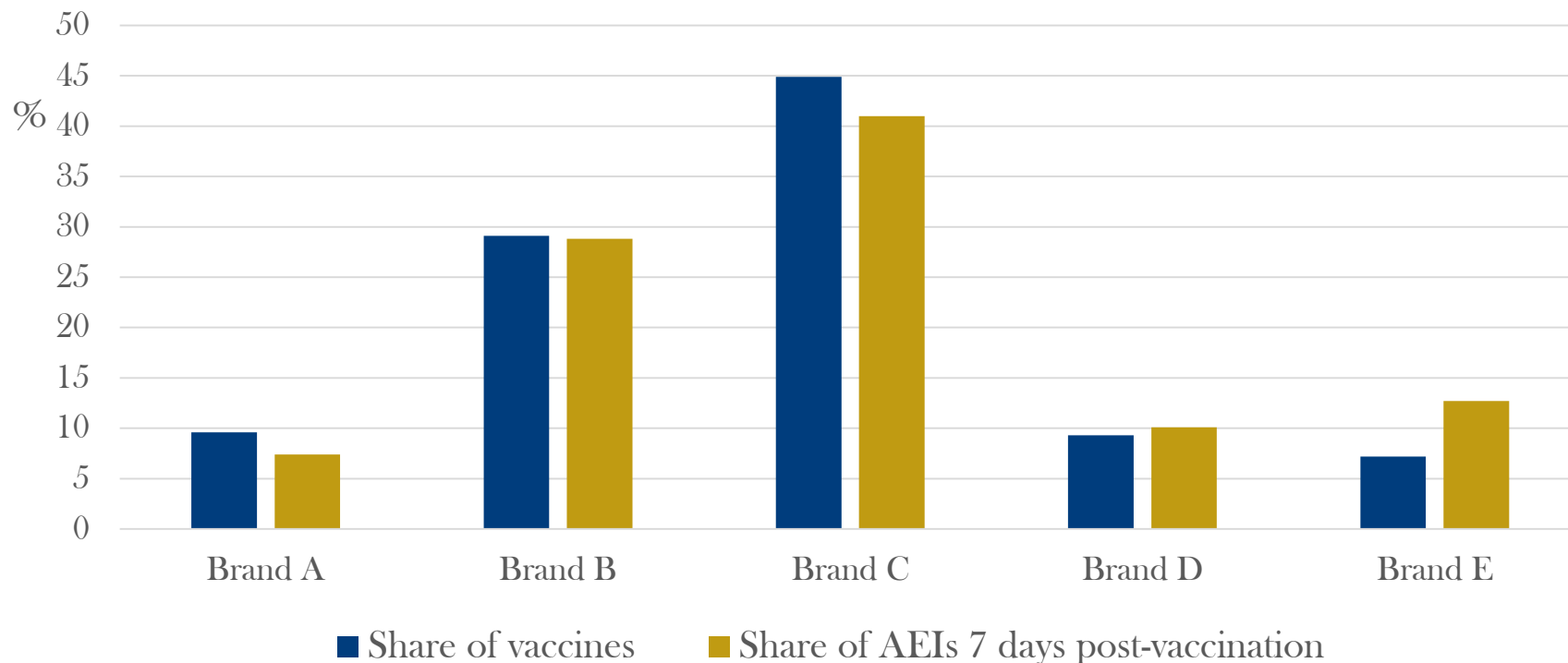
Challenge 2: Extend data
Serology

- Lancet leader called for a world serology bank

- Serological studies are traditionally either expensive or limited in the data available
  - Expensive – detailed collection using family surveys
  - Limited - surplus sera from lab tests – very limited clinical data

- We propose using sentinel networks >>

Opportunities and challenges of a World Serum Bank

The recent publication by Jessica Metcalf and others (Aug 13, p 728),[1] calls for the establishment of a World Serum Bank; something we feel should be titled a World Serology Bank given its emphasis on monitoring changes in the immune response to infections. A possible rapid and cost-effective way of setting this up would be to use existing primary care sentinel networks such as the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC).[2]

We propose inviting patients who attend their sentinel network practice for routine blood tests to also provide an additional sample for serology. These samples can be sent directly from the practice to a coordinating laboratory in the way we currently send virology and other specimens. This new approach to serology banking represents a compromise between using residual serum

Challenges: 3.
Properly validated approach

- Properly theorised approach to QI
  - Education – Audit-Based Education (ABE)
  - Behavioural change model – COM-B model

- Exemplars
  - ABE – practice specific feedback
  - ABE observational studies & trial
  - iPLAN

The COM-B system: Behaviour occurs as an interaction between three necessary conditions

Michie et al (2011) Implementation Science

Implementation Science 2009 4:39 DOI: 10.1186/1748-5908-4-39
Kidney International, 2013, 84(3) 609 DOI: 10.1038/ki.2013.96
Ann. Behav. med. (2016). DOI:10.1007/s12160-016-9843-3

Challenge 3: *Data quality*
Practice feedback – episode type
# *Audit-based education*

Table 1 compares the weekly recording rates for your practice compared with all the member practices in
the RCGP Research & Surveillance Centre. Period covered is from 27/02/2017 to 26/03/2017

Figure 1 provides a visual representation of your practice compared with the RSC as a whole.

Key conditions		Episodes of ca	
		Count	Per w
Influenza (H27%)		1	
	RSC	586	
Acute Bronchitis (H060%)		1	
	RSC	391	
Acute Bronchitis and bronchiolitis (H06%)		85	
	RSC	5993	14
Asthma (H33%)		18	
	RSC	6054	1

Table 1: How your practice recorded four key win
of care (Consultations), your practice list
ments. *rate is reported as per week per
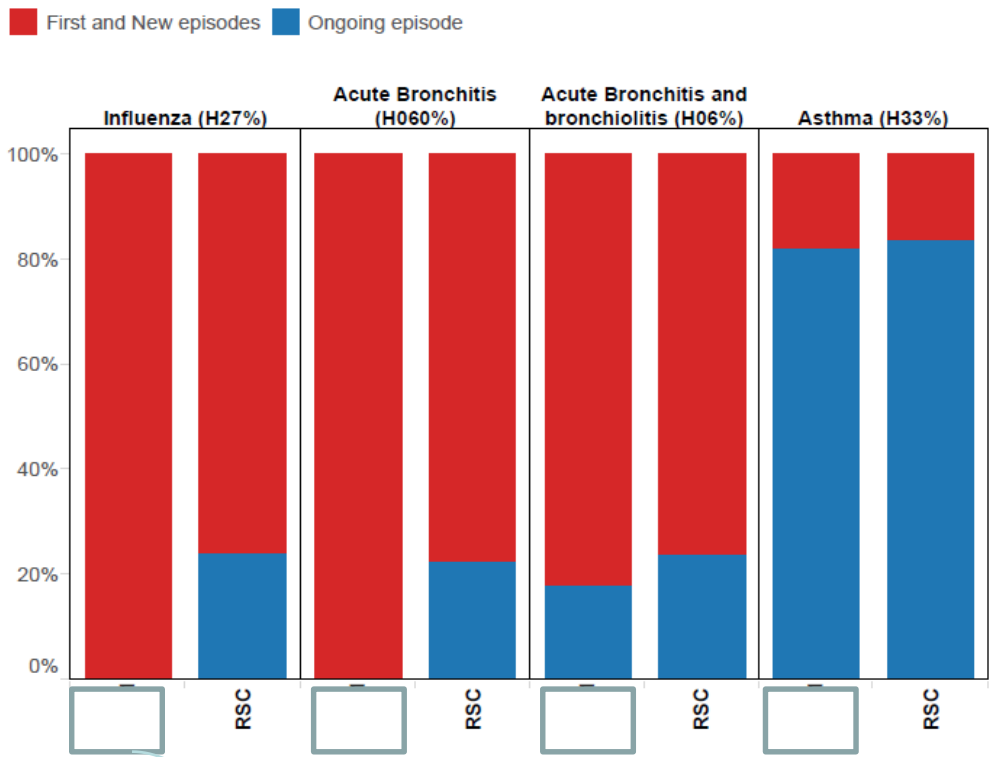
Figure 1: Your practice compared with RSC, showing the proportion of first and new to review consultation for the four marker conditions

Please use Figure 1 to compare whether you have recorded a greater or lesser first and new episode types compared with ongoing/follow-up appointments in your practice compared with the RSC as a whole. Differences in age,
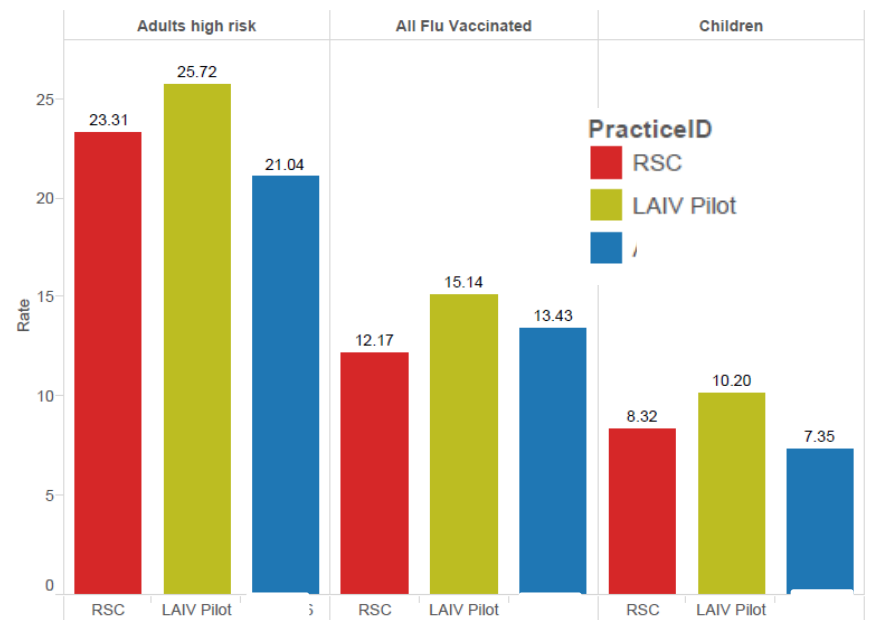
Challenge 3:
Data quality
Vaccine exposure

- **Standard feedback to practices**
  - Population
  - >65 years
  - Children (LAIV)
  - CMO high risk groups

- **LAIV pilot areas**
  - Intranasal vaccine pilot
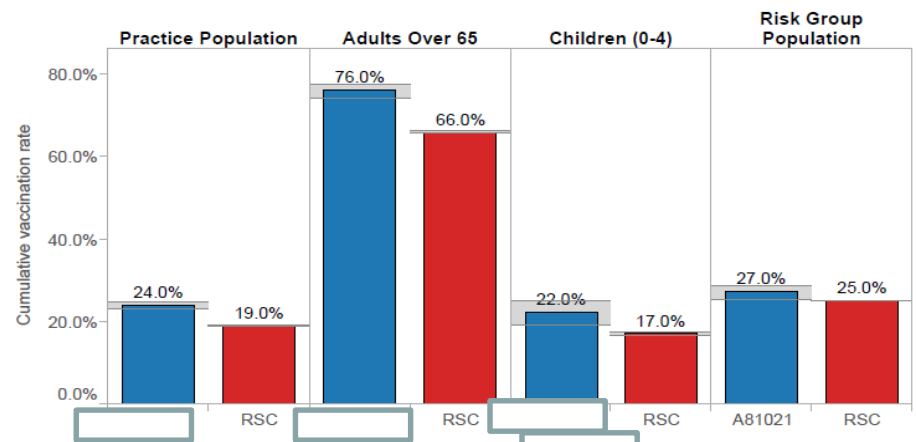
| Category	Practice ID	Cumulative vaccination rate	Cumulative vaccinated patients	Vaccinated this week patients	Vaccinated this week rate
Practice Population	A81021	24.0%	2,983	85	1.0%
	RSC	19.0%	296,410	10,749	1.0%
Adults Over 65	A81021	76.0%	1,920	42	2.0%
	RSC	66.0%	179,209	4,486	2.0%
Children (0-4)	A81021	22.0%	163	4	1.0%
	RSC	17.0%	15,603	1,254	1.0%
Risk Group Population	A81021	27.0%	651	20	1.0%
	RSC	25.0%	61,392	2,894	1.0%

Challenges: 4.
Segmentation & risk scores

1. **Few population segmentation tools**

   e.g. John Hopkins & Ambulatory Care Groups (ACGs)

2. **Mover from generic risk to modifiable**

   – e.g. Risk of Falling – using linked data
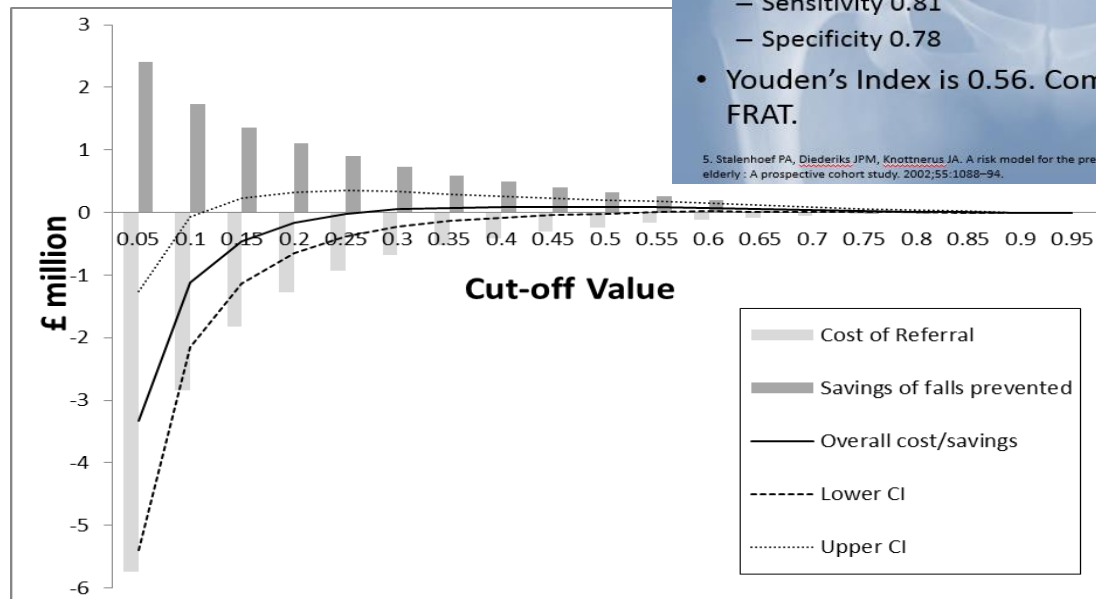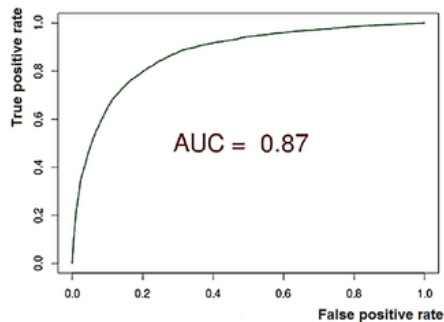
Model Performance

- AUC of the ROC curve – 0.87
  - Our previous model of GP data alone was 0.72, The next nearest performing model in literature is 0.79 (5).
- Maximum sensitivity and specificity at risk cut off of 0.07
  - Sensitivity 0.81
  - Specificity 0.78
- Youden's Index is 0.56. Compared to 0.39 for FRAT.

5. Stalenhoef PA, Diederiks JPM, Knottnerus JA. A risk model for the prediction of recurrent falls in community-dwelling elderly : A prospective cohort study. 2002;55:1088–94.

PlosOne
https://doi.org/10.1371/journal.pone.0159365

www.surrey.ac.uk

Predictor	Regression Coefficient	Odds ratio	95% CI	p Value
Age 65–69	Reference	-	-	-
Age 70–74	0.41	1.51	1.32, 1.73	<0.001
Age 75–79	1.18	3.24	2.86, 3.67	<0.001
Age 80–84	1.52	4.57	4.03, 5.17	<0.001
Age 85–89	2.00	7.37	6.46, 8.42	<0.001
Age 90–94	2.00	7.40	6.33, 8.66	<0.001
Age 95+	0.77	2.17	1.75, 2.68	<0.001
Female Sex	0.12	1.13	1.06, 1.21	<0.001
Inpatient episode previous 1 month	0.86	2.37	2.00, 2.81	<0.001
Inpatient episode previous 2 months	-0.21	0.81	0.69, 0.95	<0.01
1 outpatient visit previous 1 month	0.13	1.14	1.03, 1.26	<0.05
1–5 outpatient visits previous 12 months	0.22	1.25	1.16, 1.35	<0.001
6–10 outpatient visits previous 12 months	0.32	1.38	1.21, 1.57	<0.001
≥11 outpatient visits previous 12 months	0.36	1.43	1.20, 1.69	<0.001
A+E investigation previous 3 months	-0.15	0.86	0.74, 0.99	<0.05
Non-elective admission in previous year	0.32	1.38	1.24, 1.54	<0.001
GP code of a fracture (>6 months ago)	0.32	1.38	1.23, 1.54	<0.001
GP code of a fracture (previous 6 months)	0.95	2.58	2.04, 3.26	<0.001
Osteoporosis	0.31	1.36	1.17, 1.60	<0.001
GP/Hospital Code of fall (>6 months ago)	0.50	1.65	1.35, 2.01	<0.001
GP/Hospital code of fall (previous 6 months)	0.93	2.53	2.13, 3.01	<0.001
COPD (Long term condition)	0.20	1.22	1.09, 1.37	<0.001
History of stroke	0.16	1.18	1.06, 1.31	<0.01
Depression (Long term condition)	0.28	1.33	1.20, 1.47	<0.001
Mental Health (Long term condition)	0.39	1.48	1.20, 1.83	<0.001
Asthma (Long term condition)	0.30	1.34	1.13, 1.60	<0.01
History of urinary tract infection	0.94	2.56	2.32, 2.81	<0.001
Polypharmacy 1–4 unique drugs	0.88	2.41	2.12, 2.73	<0.001
Polypharmacy 5–9 unique drugs	1.00	2.71	2.40, 3.06	<0.001
Polypharmacy ≥10 unique drugs	1.02	2.76	2.42, 3.15	<0.001

A full list of variables considered can be found in S3 Table.

Real World Evidence (RWE) for SQIR

- Real world evidence (RWE)
- Nationally available data for RWE studies
- RCGP RSC
    - History & Role
    - Structure & Capability for linkage
    - Dataset & representativeness
    - Vaccine Effectiveness
- Quality Improvement & Research
- Conclusion:
    - There is greater capacity for collaborative research

Future collaborations

- The **RCGP RSC** network is representative of the underlying English population. It is currently under used by the academic community

- There is potential for more collaboration – findings independent of health system context

- We encourage interested researchers to attend the short courses on how to analyse primary care data offered by the university twice a year:

  - Surrey Winter Statistics School (SWISS) January 2018: Date TBA
  - Surrey Informatics Summer School (SISS) July 2018: Date TBA

- We welcome opportunities, especially for strategic collaboration:
  - For individual studies
  - To develop RWE methodologies

Thank you for listening

Simon de Lusignan

s.lusignan@surrey.ac.uk

www.surrey.ac.uk