

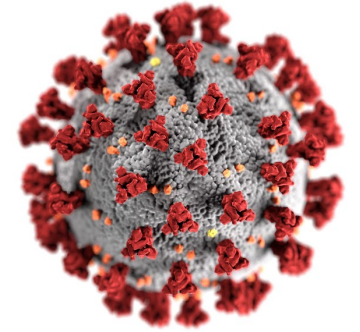
Statistics in times of increasing uncertainty

Sylvia Richardson



1 Introduction

The pandemic changed the course of my presidency ...



ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

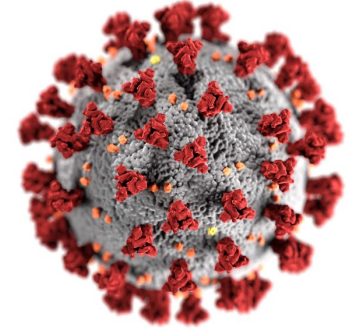
Statistics, Data and Covid

Ten statistical lessons the government can learn from the past year



1 Introduction

The pandemic changed the course of my presidency ...

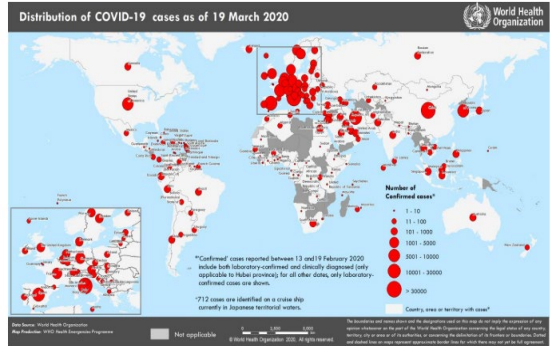


- Look back and draw out highlights of our collective engagement in the pandemic
- Discuss how the challenges faced are giving impetus to new directions
- Present our strategic thinking regarding our engagement in data science.



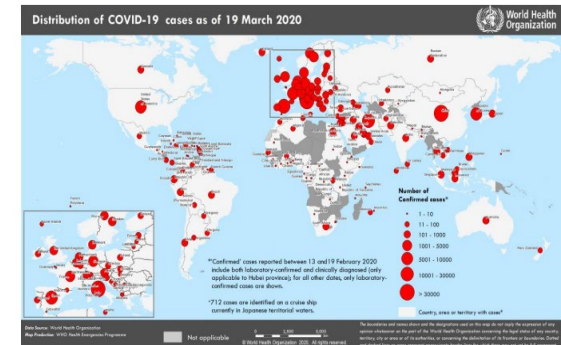
2 Statistics making an impact/ Rules of engagement

- Statistics is about matters of the highest importance in human affairs (Pullinger, 2013)



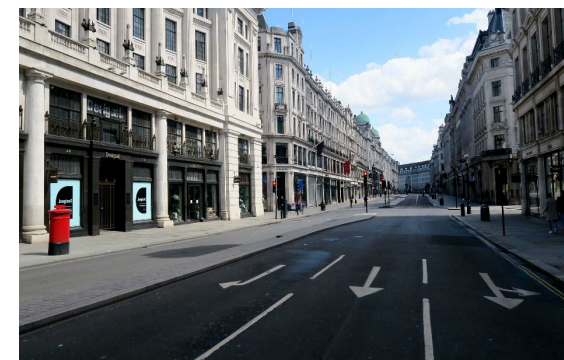
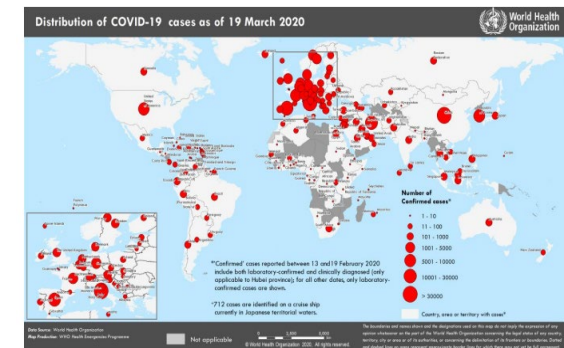
2 Statistics making an impact/ Rules of engagement

- Statistics is about matters of the highest importance in human affairs (Pullinger, 2013)
- Our 4 strategic objectives have guided our involvement in the pandemic
 - for statistics to be used effectively in the public interest
 - for statistics as a discipline to thrive
 - for a strong body of professional statisticians
 - for society to be more statistically literate.



2 Statistics making an impact/ Rules of engagement

- Statistics is about matters of the highest importance in human affairs (Pullinger, 2013)
- Our 4 strategic objectives have guided our involvement in the pandemic
- Ingredients for impact
 - Willingness to engage
 - Statistically principled reasoning
 - Focussing on statistical issues and communication.



3 Referring to statistical principle: our shared modus operandi



3 Referring to statistical principle: our shared modus operandi

Questioning the “value” of different data sets



3 Referring to statistical principle: our shared modus operandi

Questioning the “value” of different data sets



Vaccinations			
People vaccinated in England	Last 7 days – first dose	Last 7 days – second dose	Last 7 days – dose
NATION	7,505	19,125	16,173
<small>Up to and including 7 September 2022</small>			
All vaccinations data in England	Total – first dose 45,243,079	Total – second dose 42,653,942	Total – boosters 33,534,600

Cases	Deaths	Healthcare
People tested positive in England	Deaths within 28 days of positive test in England	Patients in hospital in England
NATION	NATION	NATION
<small>Up to and including 3 September 2022</small>		
Last 7 days 24,880 ↑ 131 (0.5%)	Last 7 days 384 ↓ -127 (-24.9%)	Last 7 days 3,628 ↓
	<small>▶ Rate per 100,000 people: 0.7</small>	



3 Referring to statistical principle: our shared modus operandi

Questioning the “value” of different data sets

- Causal link between symptoms and swab testing: leads to bias in raw prevalence estimates



Vaccinations			
People vaccinated in England	Last 7 days – first dose	Last 7 days – second dose	Last 7 days – dose
NATION	7,505	19,125	16,173
Up to and including 7 September 2022			
All vaccinations data in England	Total – first dose	Total – second dose	Total – boost
	45,243,079	42,653,942	33,534,6
Cases			
People tested positive in England	Deaths		Healthcare
NATION	Deaths within 28 days of positive test in England		Patients a
Up to and including 3 September 2022	NATION		NATION
	Up to and including 3 September 2022		Up to and includ
Last 7 days	Last 7 days	Last 7 days	
24,880 ↑ 131 (0.5%)	384 ↓ -127 (-24.9%)	3,628 ↓	
	▶ Rate per 100,000 people: 0.7		

3 Referring to statistical principle: our shared modus operandi

Questioning the “value” of different data sets

- Causal link between symptoms and swab testing: leads to bias in raw prevalence estimates
- Surveys based on random sampling: Imperial REACT study and the ONS Infection Survey



Vaccinations			
People vaccinated in England			
	Last 7 days – first dose	Last 7 days – second dose	Last 7 days – dose
NATION	7,505	19,125	16,173
Up to and including 7 September 2022			
All vaccinations data in England			
	Total – first dose	Total – second dose	Total – boosters
	45,243,079	42,653,942	33,534,614

Cases	Deaths	Healthcare
People tested positive in England	Deaths within 28 days of positive test in England	Patients in hospital in England
NATION	NATION	NATION
Up to and including 3 September 2022		
Last 7 days 24,880 ↑ 131 (0.5%)	Last 7 days 384 ↓ -127 (-24.9%)	Last 7 days 3,628 ↓
	▶ Rate per 100,000 people: 0.7	

Imperial College
London

REal-time Assessment
of Community
Transmission

 Office for
National Statistics

Coronavirus (COVID-19) Infection Survey, UK: 2
September 2022



3 Referring to statistical principle: our shared modus operandi

Questioning the “value” of different data sets

- Causal link between symptoms and swab testing: leads to bias in raw prevalence estimates
 - Surveys based on random sampling: Imperial REACT study and the ONS Infection Survey
- Statistical conversation on different forms and means to ramping up active surveillance.



Vaccinations			
People vaccinated in England			
	Last 7 days – first dose	Last 7 days – second dose	Last 7 days – dose
NATION	7,505	19,125	16,173
<small>Up to and including 7 September 2022</small>			
All vaccinations data in England			
	Total – first dose	Total – second dose	Total – boosters
	45,243,079	42,653,942	33,534,600
Cases			
People tested positive in England			
NATION			
<small>Up to and including 3 September 2022</small>			
Last 7 days	Last 7 days		
24,880	↑ 131 (0.5%)	384	↓ -127 (-24.9%)
<small>▶ Rate per 100,000 people: 0.7</small>			
Deaths			
Deaths within 28 days of positive test in England			
NATION			
<small>Up to and including 3 September 2022</small>			
Healthcare			
Patients in hospital in England			
NATION			
<small>Up to and including 3 September 2022</small>			
Last 7 days	Last 7 days		
3,628	↓ -127 (-24.9%)		

Imperial College
London

REal-time Assessment
of Community
Transmission

 Office for
National Statistics

Coronavirus (COVID-19) Infection Survey, UK: 2
September 2022



3 Referring to statistical principle: our shared modus operandi

Adding value through evidence synthesis

- System's view of disease surveillance benefits from data synthesis of different sources
- Statistical fusion of Test & Trace testing data with REACT

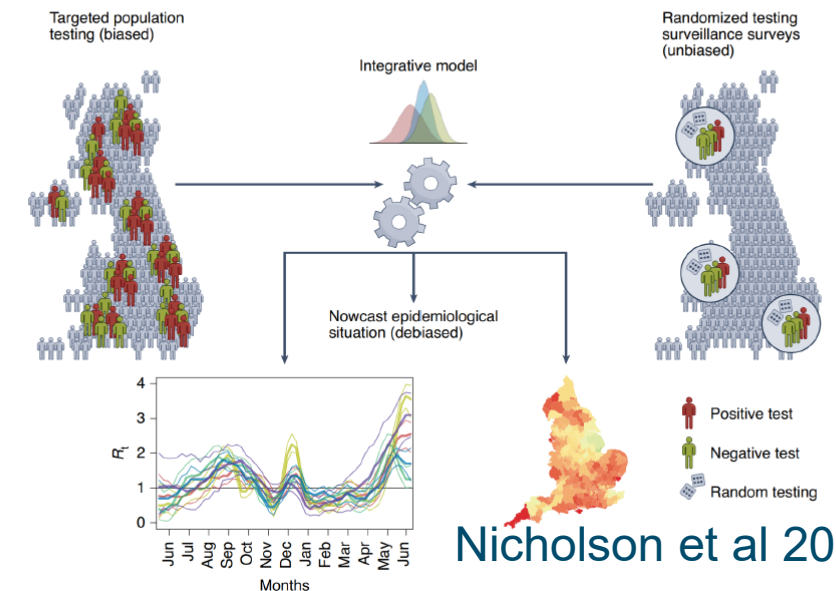


3 Referring to statistical principle: our shared modus operandi

Adding value through evidence synthesis

- System's view of disease surveillance benefits from data synthesis of different sources
 - Statistical fusion of Test & Trace testing data with REACT
- weekly debiased estimate of true prevalence at a local scale

Randomized surveillance – essential for understanding bias

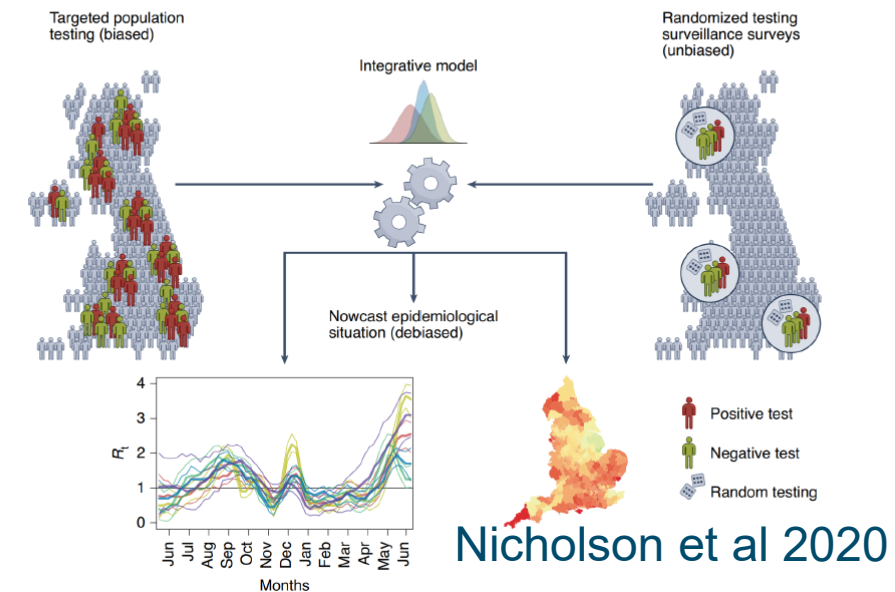


3 Referring to statistical principle: our shared modus operandi

Adding value through evidence synthesis

- System's view of disease surveillance benefits from data synthesis of different sources
 - Statistical fusion of Test & Trace testing data with REACT
- substantial challenges in operationalising data synthesis at pace.

Randomized surveillance – essential for understanding bias



3 Referring to statistical principle: our shared modus operandi

Evaluation of treatments, health surveillance systems and policies

- Can we strive to learn at the same time as acting ?



3 Referring to statistical principle: our shared modus operandi

Evaluation of treatments, health surveillance systems and policies

- Platform trials: **RECOVERY** success stems from conjunction of trial design community & responsive regulatory framework & NHS



3 Referring to statistical principle: our shared modus operandi

Evaluation of treatments, health surveillance systems and policies

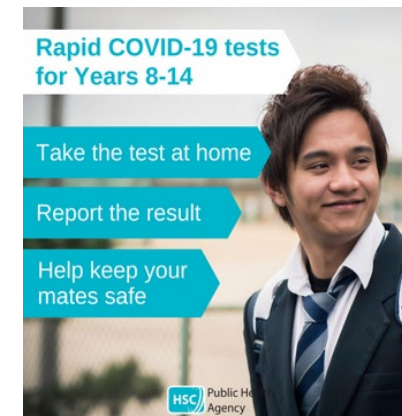
- Platform trials: **RECOVERY** success stems from conjunction of trial design community & responsive regulatory framework & NHS
- Evaluation of new systems and policies?
 - Piecemeal rather than planned from the start
 - Challenging to target relevant policies in a fast moving pandemic.



Test and Trace Business Plan

Helping to break chains of COVID-19 transmission, protect the public's health, and enable people to return towards a more normal way of life: The next phase of NHS Test and Trace

Published 10 December 2020



4 Challenges

Communication

- The wider public was continually assailed by covid related statistics difficult to interpret



The Guardian
Zero': how the UK papers covered a day without a ...

4 Challenges

Communication

- The wider public was continually assailed by covid related statistics difficult to interpret
- The membership, RSS ambassadors and the Task Force engaged in sustained efforts of communication

Covid-19 FAQs

The coronavirus pandemic has led to many different statistics being presented to us from the media, official sources and perhaps more unofficial sources. But what do they all mean, how should we interpret them, and in what ways do they affect us personally?

What questions should you ask when you hear a claim based on data?

David Spiegelhalter and Anthony Masters

The source, the number, and the claim need to be trustworthy

Jen Rodgers' podcasts



4 Challenges

Communication

- The wider public was continually assailed by covid related statistics difficult to interpret
- The membership, RSS ambassadors and the Task Force engaged in sustained efforts of communication

→ Championing impartial evidence is a skill to learn and enjoy.

What questions should you ask when you hear a claim based on data?

David Spiegelhalter and Anthony Masters

The source, the number, and the claim need to be trustworthy

Jen Rodgers podcasts



4 Challenges

Communicating structural and statistical uncertainty

- an arduous but necessary task to inform the public debate and counteract misinformation
- open scrutiny of the plausibility of assumptions

*“La vraie science est une ignorance qui se sait”
(Michel de Montaigne)*



4 Challenges

Communicating structural and statistical uncertainty

- an arduous but necessary task to inform the public debate and counteract misinformation
 - foster good practice in the media
- build on our experience to support good presentation of evidence for all future societal issues.

*“La vraie science est une ignorance qui se sait”
(Michel de Montaigne)*



4 Challenges

Agility in policy evaluations

- Integrating insights from simulation exercises, e.g. agent-based models
 - qualitative and quantitative understanding of key drivers
 - feed into policy trials

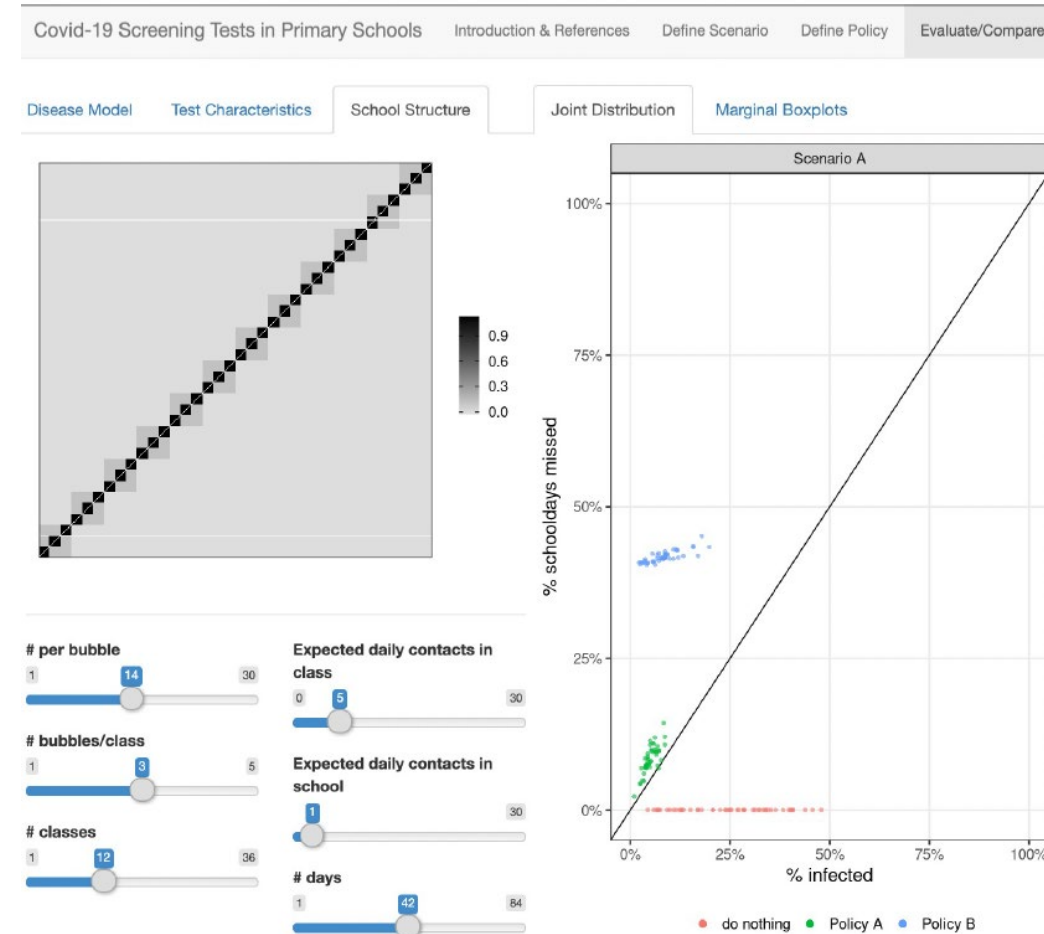


4 Challenges

Agility in policy evaluations

- Integrating insights from simulation exercises, e.g. agent-based models
 - qualitative and quantitative understanding of key drivers
 - feed into policy trials

Agent based modelling to compare school policies for controlling covid-19 transmission



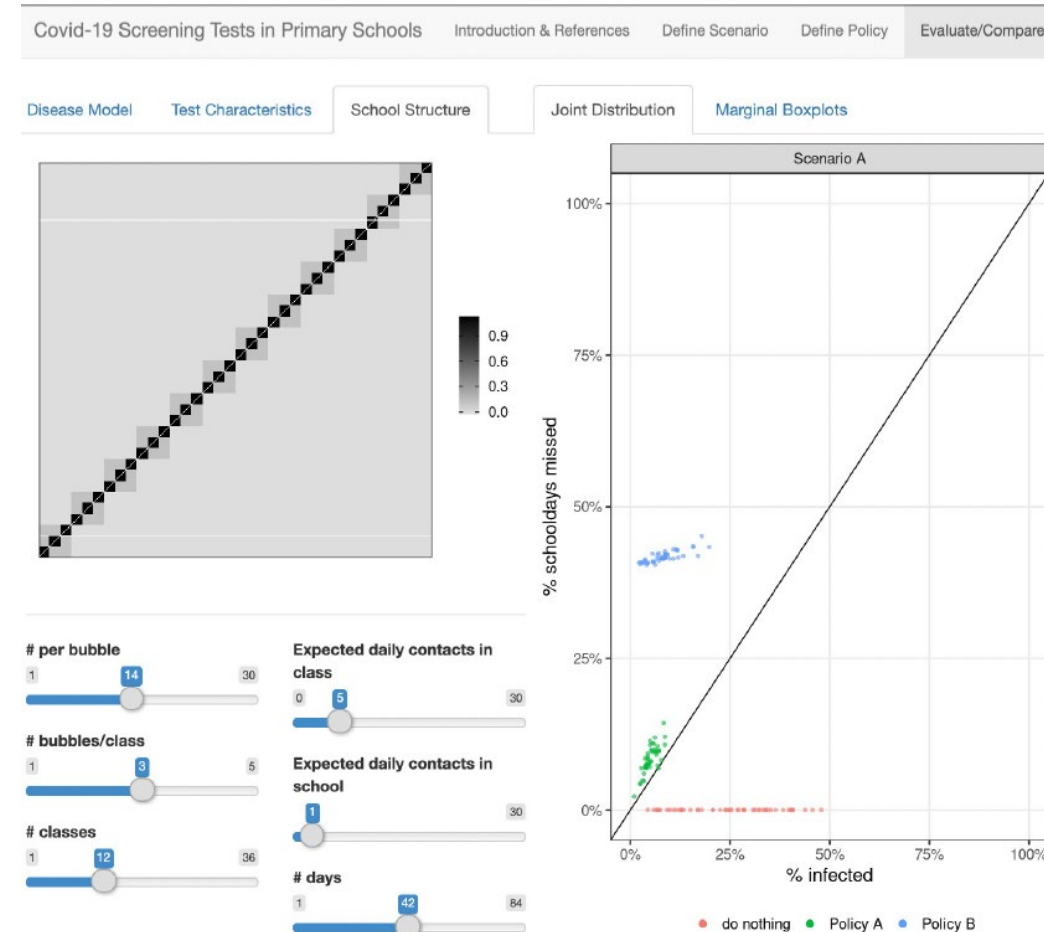
Kunzmann et al 2022

4 Challenges

Agility in policy evaluations

- Integrating insights from simulation exercises, e.g. agent-based models
 - qualitative and quantitative understanding of key drivers
 - feed into policy trials
- Adaptive strategies for *merging* simulation-based and designed evaluations is a promising avenue.

Agent based modelling to compare school policies for controlling covid-19 transmission



Kunzmann et al 2022

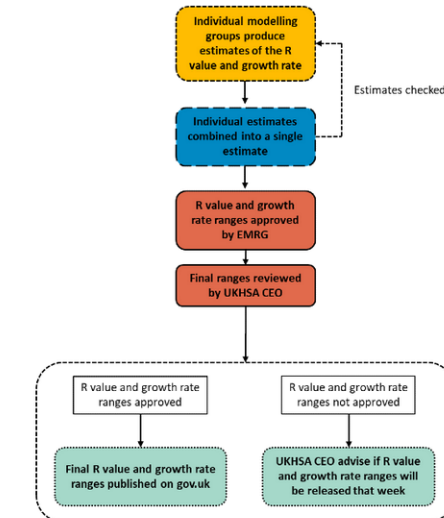
4 Challenges

The statistician as expert

- Distinction between *presenting evidence* and *its consequences* is easily blurred
- Contribution of statistical expertise to public health decisions mainly through SPI-M + other data driven initiatives

Methodology used by SPI-M for estimating R and growth rate

Flowchart: estimating the R value and growth rate – from production to publication



16 September 2022 — Guidance
The R value and growth rate

4 Challenges

The statistician as expert

- Distinction between *presenting* evidence and its *consequences* is easily blurred
- Contribution of statistical expertise to public health decisions mainly through SPI-M + other data driven initiatives
- How statisticians and data scientists could and – should – be involved as experts in situation of emergency in the future?

→ train a reservoir of statisticians.

Covid evidence sessions

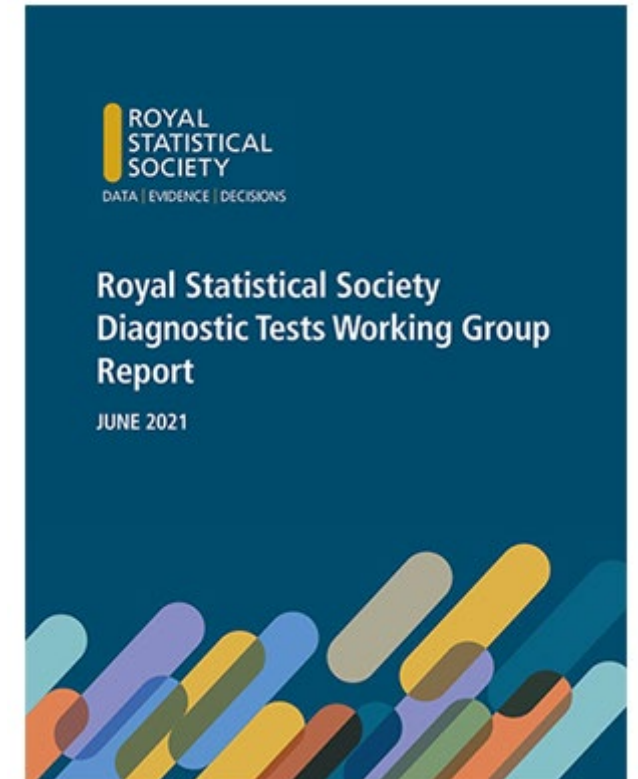
Watch the third session, *Evidence and policy-making*



4 Challenges

Road to long-term impact of recommendations is long ...

- Focus on topics where there is concern that issues of statistical nature are ignored
 - diagnostic tests
- Create the basis of a dialog with responsive health authorities
 - Working Group and Report (2021)
 - Recommendations on standards for study design, presentation of evidence, planning for future pandemics, regulations.



5 Statistical legacy from the pandemic

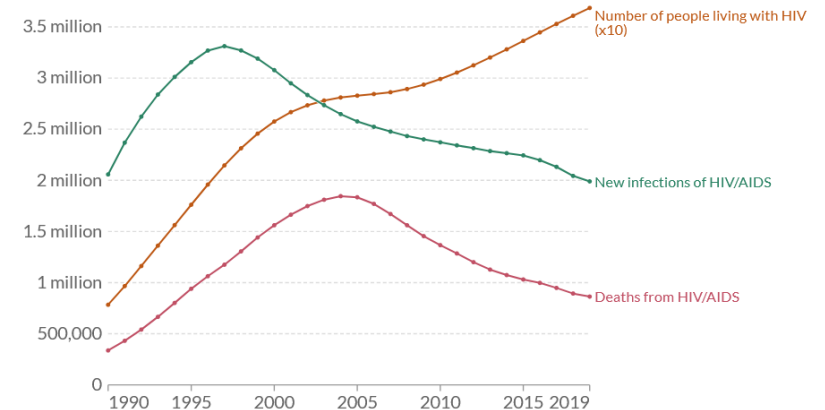
Throughout history, advances in statistics have arisen from demands in science and in society

- Before Covid-19 last global infectious disease was HIV/AIDS epidemic
 - led to substantial advances in survival analysis and partially observed processes.
- New directions created by the specific context of the SARS-CoV-2 pandemic?

Prevalence, new cases and deaths from HIV/AIDS, World, 1990 to 2019

To fit all three measures on the same visualization the total number of people living with HIV has been divided by ten (i.e. in 2019 there were 36.8 million people living with HIV).

↔ Change country



Source: IHME, Global Burden of Disease (2019)

▶ 1990 ◯ 2019

5 Statistical legacy from the pandemic

*Modelling and inference **at pace***

- Fitting of complex models to imperfect, incomplete and sometimes conflicting data sources
 - Sensitivity to model form and noise structure
 - Approximate computations
 - Interpretation /(mis) of different metrics
 - Back loop between behavioural changes and epidemic evolution
 - Evaluation.

RSS call out
Special Topic meeting
on
R/localR/transmission
of Covid-19

R-VALUE 

5 Statistical legacy from the pandemic

Modelling and inference at pace

- Fitting of complex models to imperfect, incomplete and sometimes conflicting data sources
- Continuing a reflection on the impact of the intensive modelling work with a view to drawing lessons for improved preparedness and communication.

RSS call out
Special Topic meeting
on
R/localR/transmission
of Covid-19

R-VALUE 

5 Statistical legacy from the pandemic

Statistical interoperability

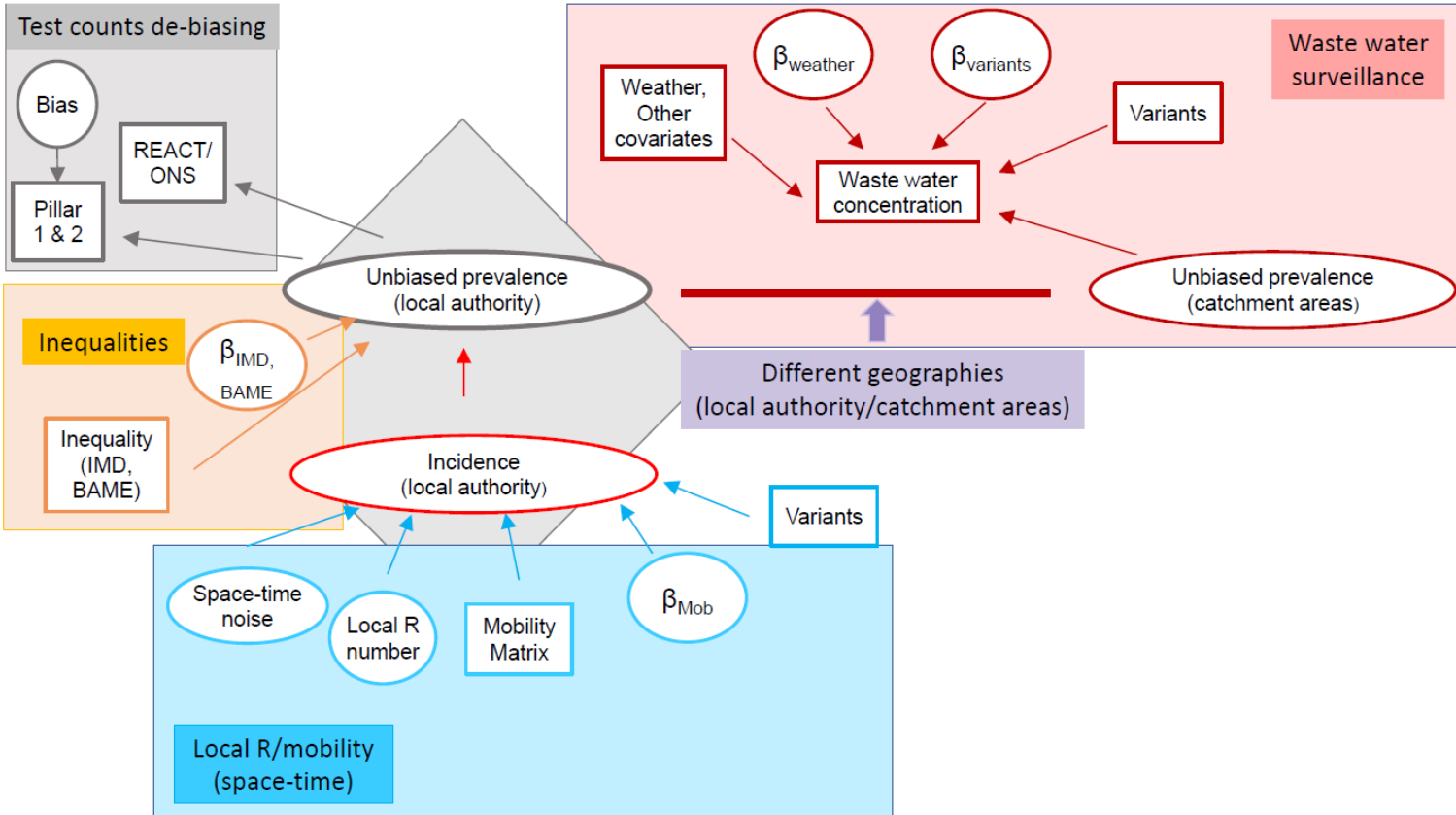
- Agility in statistical modelling, while preserving coherent treatment of uncertainty
- Practice developed by the Turing-RSS Health Data Lab
 - **Agility:** modular inference where outputs become inputs
 - **Sustainability:** common open source code
 - **Transferability:** co-ownership of projects
 - **Preparedness:** quick repurposing of solutions.



5 Statistical legacy from the pandemic

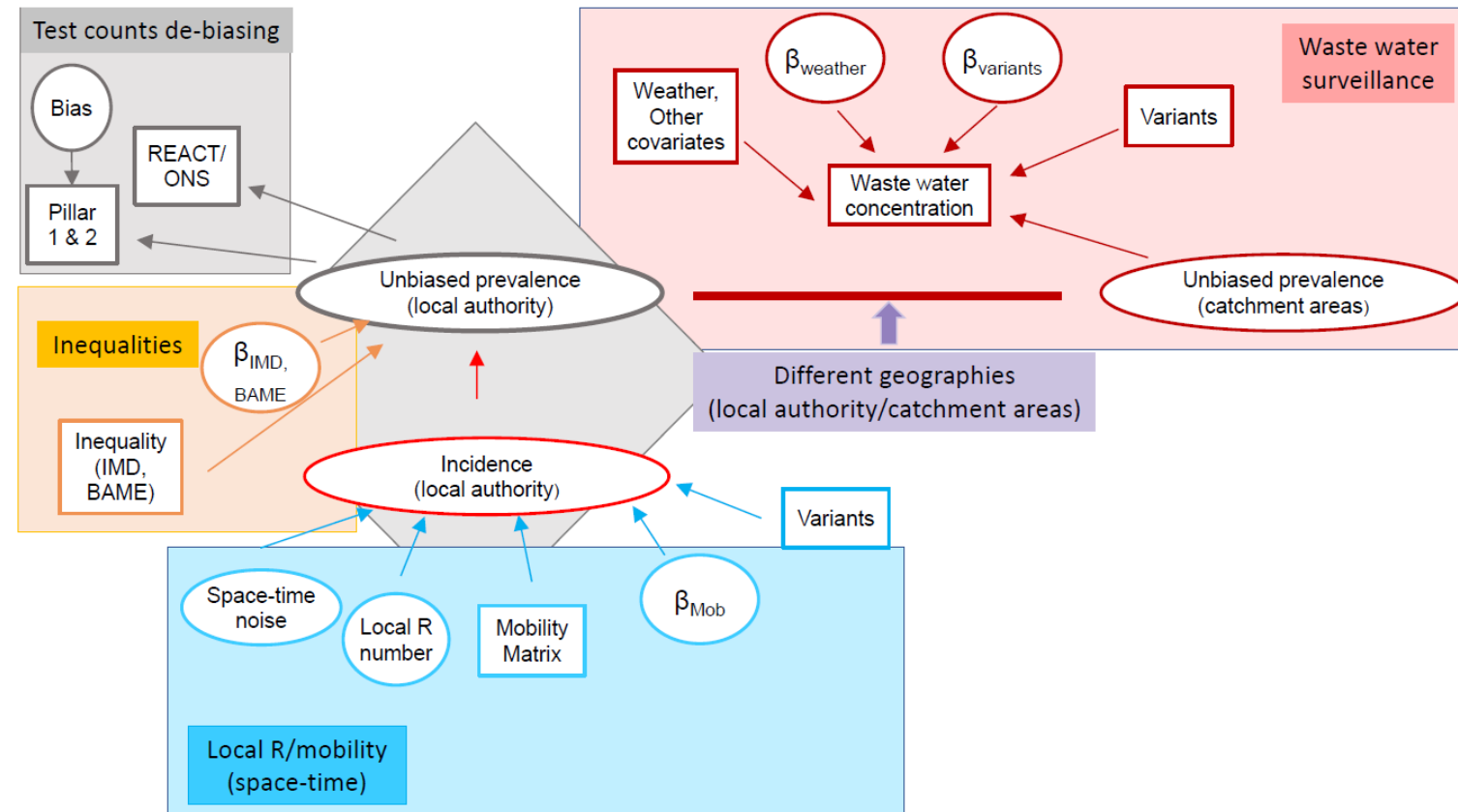
Statistical interoperability

Schematic graphical representation of interoperability between projects undertaken by the Turing-RSS Health Data Lab



5 Statistical legacy from the pandemic

→ Progress our knowledge and practice of efficient and principles transfer of information and uncertainty between submodels.



6 Changing world of data driven activities

Changing data landscape

- Questioning of the place of statistics
- Recurrent theme in previous addresses (Hand, Diggle, Ashby)

RSS Data Science Task Force

- Overall strategy of visible engagement with DS
- Three major areas of investment.



6 Changing world of data driven activities

Three major areas of investment

- Alliance for Data Science Professionals:
 - fostering a highly skilled DS workforce

The Alliance for Data Science Professionals



Supported by:

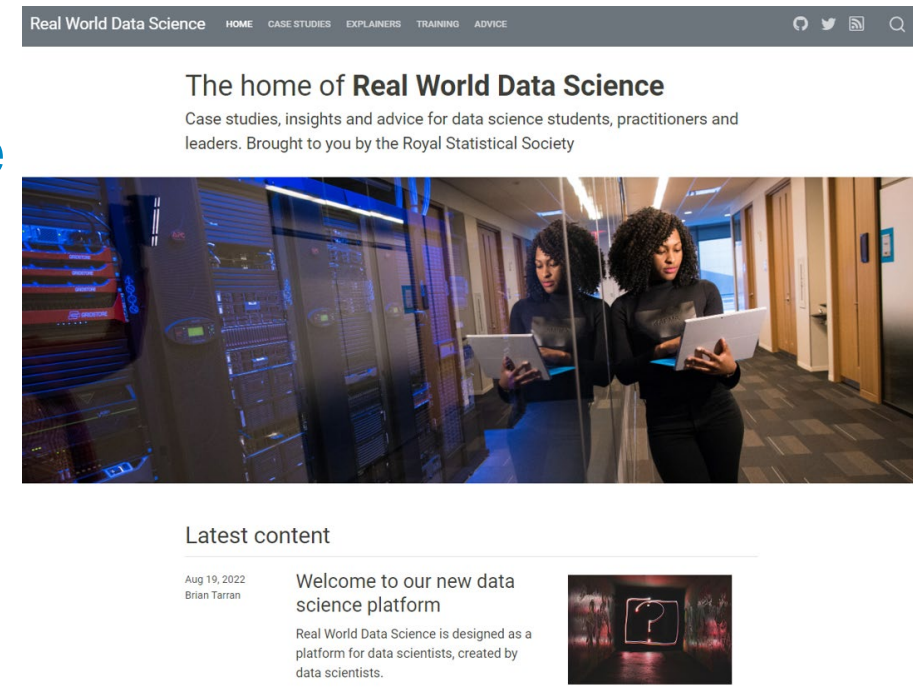
THE ROYAL SOCIETY



6 Changing world of data driven activities

Three major areas of investment

- Alliance for Data Science Professionals:
 - fostering a highly skilled DS workforce
- New RSS Real World Data Science platform:
 - resources of practical benefit to DS practitioners



The screenshot shows the homepage of the Real World Data Science website. The header includes the site name and navigation links: HOME, CASE STUDIES, EXPLAINERS, TRAINING, and ADVICE. The main heading is "The home of Real World Data Science", followed by a sub-heading: "Case studies, insights and advice for data science students, practitioners and leaders. Brought to you by the Royal Statistical Society". Below this is a large image of two women in a server room looking at tablets. A "Latest content" section features a post from August 19, 2022, by Brian Tarran, titled "Welcome to our new data science platform". The post text states: "Real World Data Science is designed as a platform for data scientists, created by data scientists." A small image of a question mark is next to the text.

6 Changing world of data driven activities

Three major areas of investment

- Alliance for Data Science Professionals:
 - fostering a highly skilled DS workforce
- New RSS Real World Data Science platform:
 - resources of practical benefit to DS practitioners
- New online open access data science journal
- pan data science ambition, cross fertilisation and unifying themes.



7 Concluding remarks

Lessons drawn

- on principles of engagement,
- role of statisticians faced with emergency,
- capacity for impact.

Entering a period of major world uncertainties

- focus on statistical agility to tackle substantial uncertainties,
- pivotal strategic engagement towards data science.



7 Concluding remarks

We all miss Sir David Cox's insightful voice

“In principle, decisive actions can be combined with intellectual appreciations that there are uncertainties in the key evidence in which decision making is based. In theoretical terms these are the dual themes of decision and inference that reappear throughout recent discussions of general principles in statistics.

The possibility of explicit and quantitative resolution of this conflict is one of the most important intellectual contributions to our subject, with far-reaching and as yet undeveloped applications” (Cox, 1981).

