Digital Health Surveillance

Unlocking the potential for real-time health data analysis in the Maltese Islands

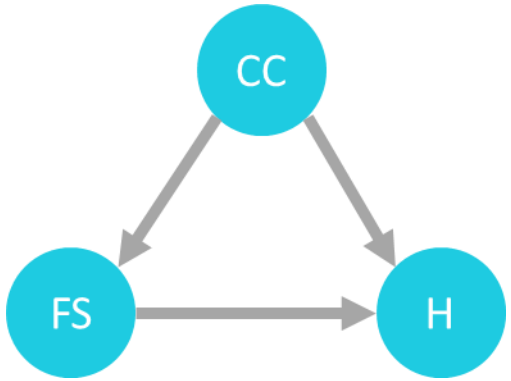
Queen Mary University of London
Malta Campus

Dr John Paul Cauchi MD, MSc, PhD

Educational journey

- Studied Medicine at the University of Malta
- Specialised at the LSHTM in Environmental public health
- Worked at UoM on ImaGenX, looking at environmental and lifestyle risk factors of Breast Cancer in Malta
- PhD in Brisbane, Australia, focusing on Climate Change, Food Security and Health. Fieldwork in Kiribati.

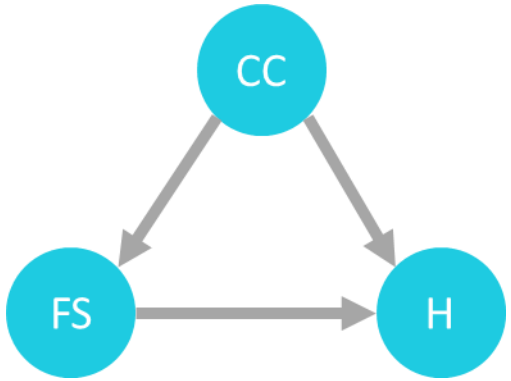
Research



- Focus: Climate change, Food security and Health interactions
- Mixed methods, applying both qualitative and quantitative data analysis to triangulate closer to the truth.
- This experience taught me the importance of accurate data collection, ascertaining meaning from numbers and words, and applying that information.

- Data was often missing, making it very hard for me to come to conclusions

Research



- Data which is absent can be both disappointing and dangerous – it fails to give you the complete picture, and makes it difficult to know the present situation.

- **Data is key**

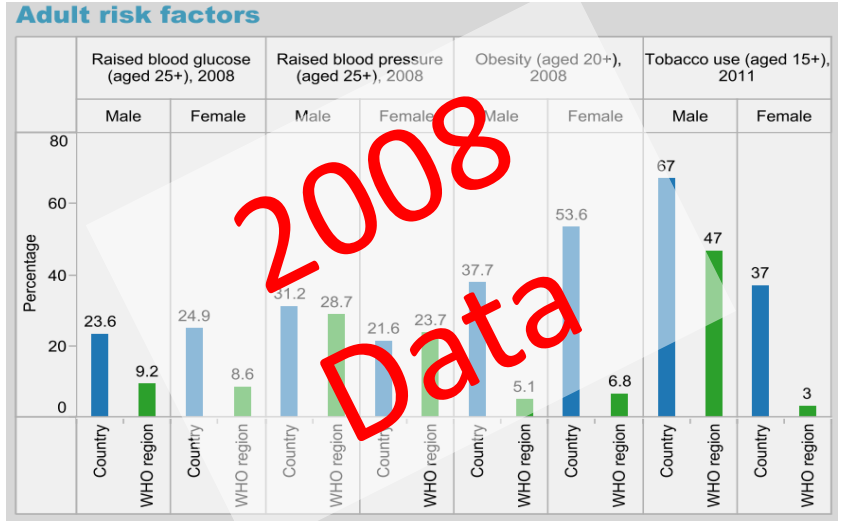
- **Data analysis** is a process of inspecting, **cleaning, transforming, and modeling** data with the goal of **discovering useful information**, **informing conclusions**, and supporting **decision-making.**

Attempted to get datasets for analysis in the form of Time Series Analysis for correlations:

- Precipitation
- Temperature
- Health data

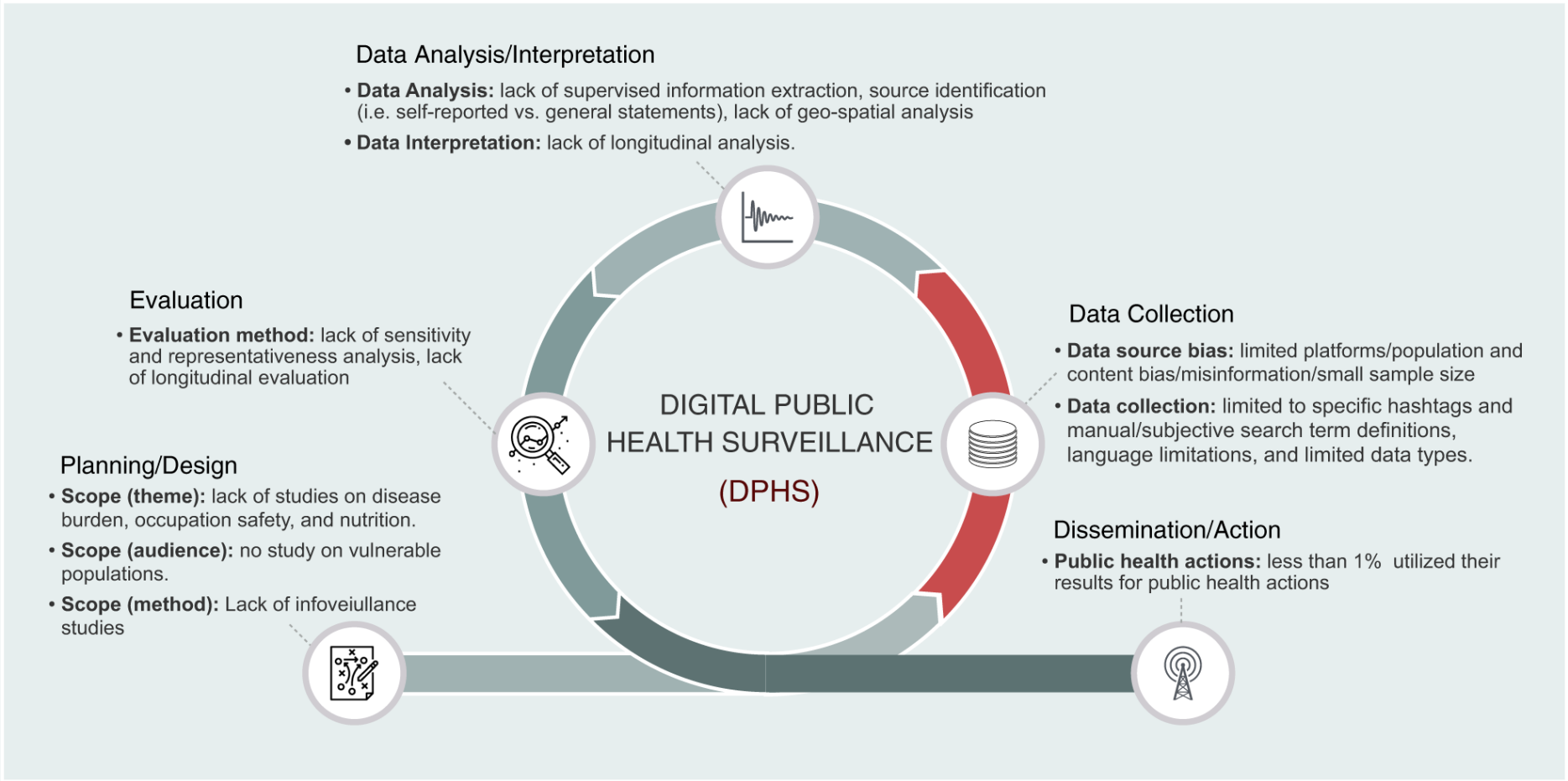
Significant gaps in data

- *E.g. dataset obtained for monthly rainfall in Nonouti from 1953 to 2013 had Jan 1991 to Sep 2010 missing…*




How to transform this data into something useful?

Public Health Data Surveillance



Public Health Data Surveillance

- 2020 – ongoing: Public Health IDCU department (currently part-time)
  - Working on Malta's Digital Health surveillance.
  - In 2020, the ECDC basically complained that John Hopkin's University (left) in the US had better data on COVID-19 in Europe than the ECDC did!
  - E-SARI Network was born – involving 15 countries that pool their data on Severe Acute Respiratory Illness (SARI), which can be used for both decisions in policy-making AND in ascertaining vaccine effectiveness in Europe.

Public Health Data Surveillance

- Problems identified:
  - Poor data collection methodologies
  - Lack of standardisation
  - Different definitions for the same illness
  - Archaic methods of data collection (Malta included)
  - Lack of coordination
  - Poor stakeholder engagement

Malta had significant problems

- Malta had only JUST shifted from a paper-based system of reporting health data to a digital one!

Patient enters admission → History taken → Physical paper files

Advantages: Standard format

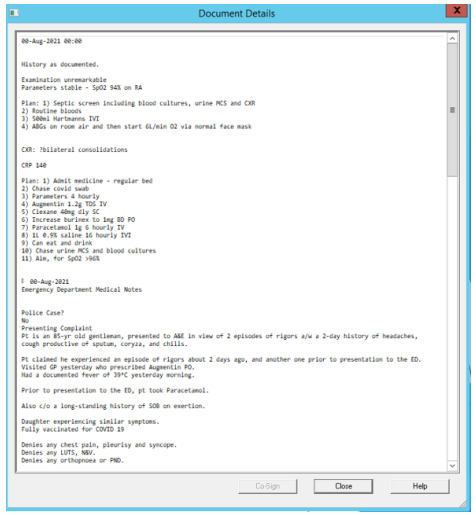
Disadvantages: Difficult to access, cumbersome, error prone

Impossible to carry out data surveillance

New system – step in the right direction, but far from ideal…

- The new system was… a step in the right direction. However…

Patient enters admission → History taken → Electronic paper files



Covid-19 accelerated an effort to digitise the system of data entry at admissions

Data is now entered onto a system as text. Effectively an ‘electronic paper’. At least it is accessible from computers!

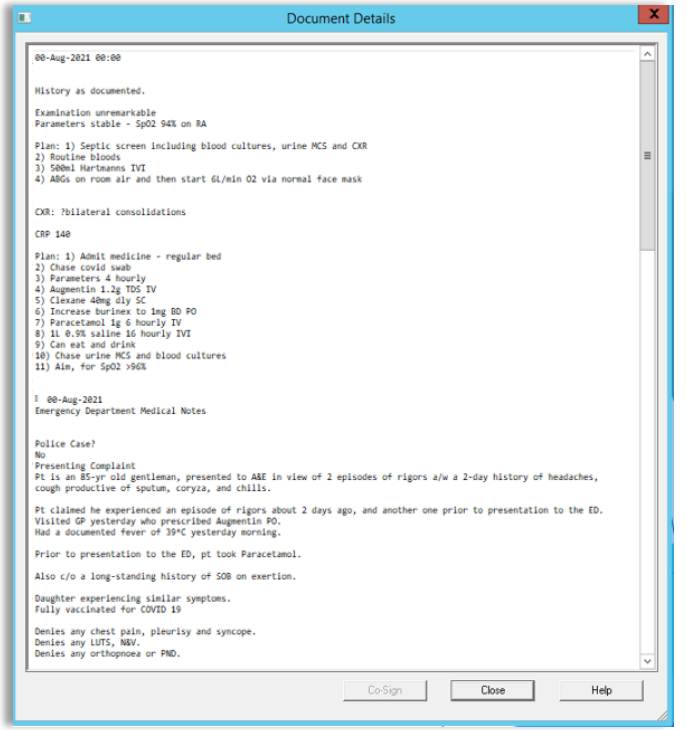
Still insufficient. More easily accessible, but no means of extracting data other than manual extraction

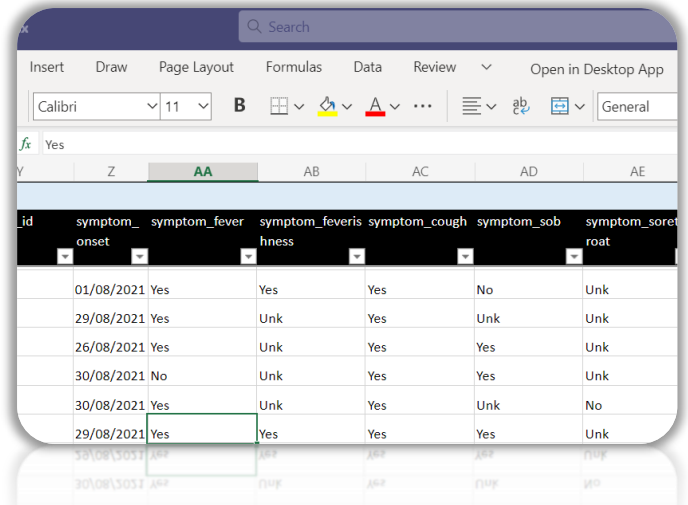
A significant change, but long way ahead

Questionnaire based surveillance

- Data extraction of **symptoms and comorbidities** commences using available software

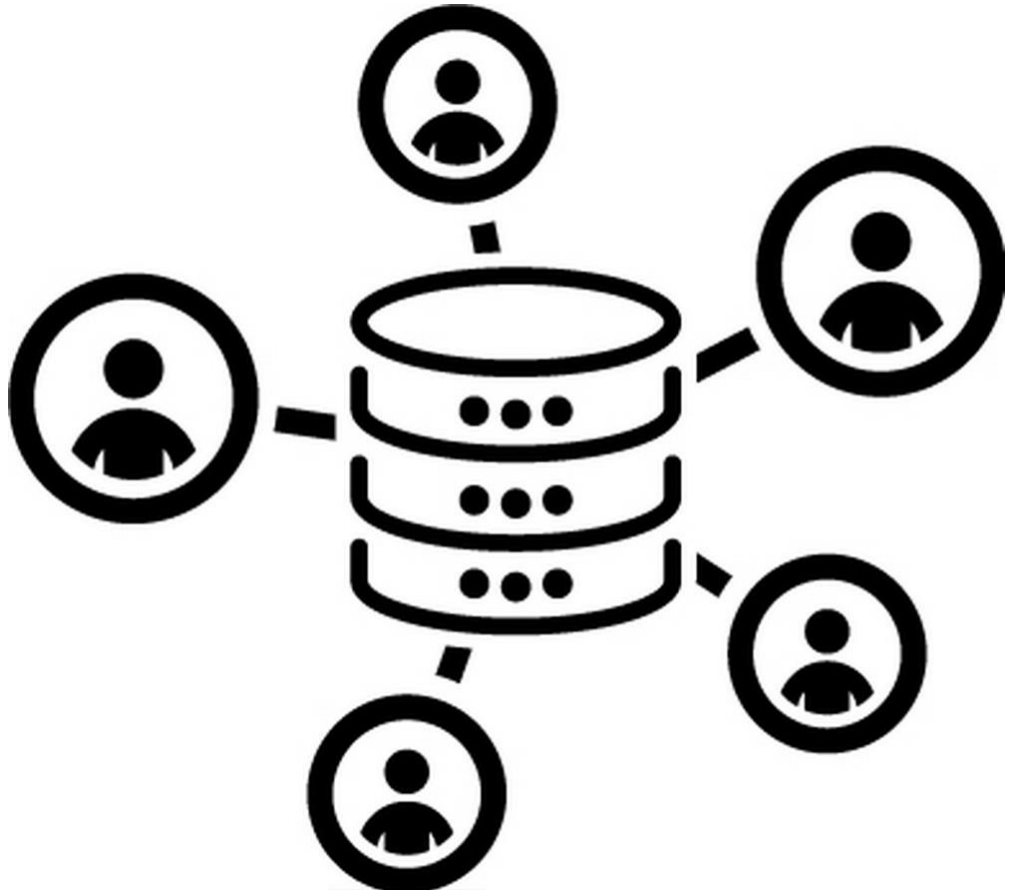
- Manual extraction is
  - **time-consuming**
  - **prone to human error** both at data entry and data extraction level





Over time, we identified other sources of data

- COVID-19 testing
- Respiratory panel of 32 pathogens
- COVID-19 vaccinations database
- Mortality data
- Hospital census
- Influenza vaccination



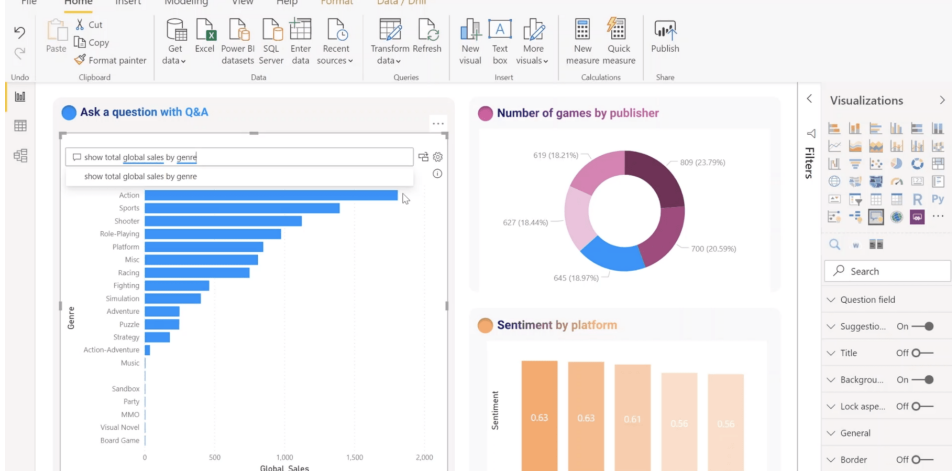
Tools required

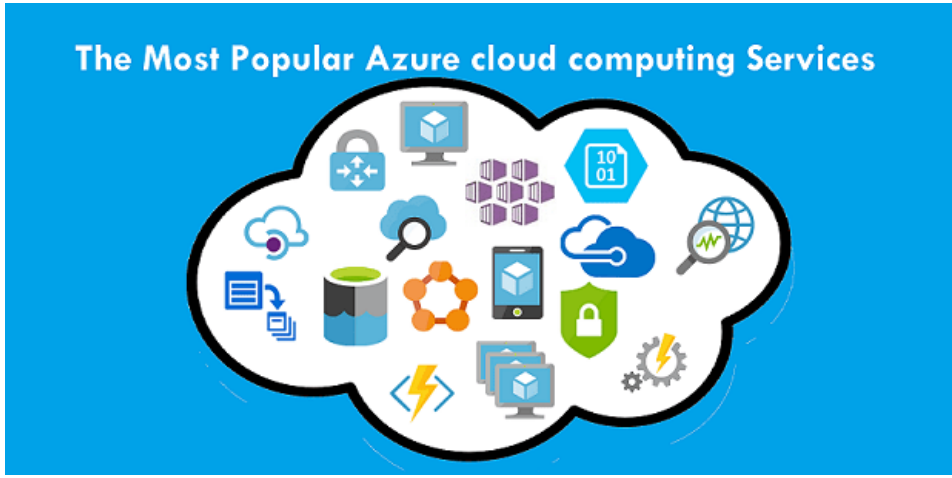
- **R software – statistical computing**
- This software enabled us to link data from various databases using ID as a common identifier.
- Efficient, constantly upgradeable, allows heavy statistical analysis
- Magnitudes of speed better than using colour-coded excel sheets for work processing.

Other Software being explored

- **Microsoft Azure Text Analytics**
- **Power BI**
- **SQL**

- These software packages are being explored for their potential to extract data automatically, or to visualise data in a compelling way.





Patient Dashboard: mid 2023

- This system would involve **entering of symptoms into fields** that can be extracted from a database, allowing automation
- **Cases can be extracted based on symptoms entered into respiratory forms**
- Could also be **expanded to other illnesses**, allowing real-time surveillance
- This would permit faster, better decision-making. It could also provide real-time information of outbreaks based on symptom combinations.

The importance of Health Data Analysis

- Can be integrated with other systems, such as Environmental Data, Zoonotic data which could lead to expanded surveillance.
- Allows for public health research to be carried out locally using large datasets.
- This can improve Malta's public health capacity, audit its abilities and improve on what we already know.
- Helps us be prepared for the next pandemic!

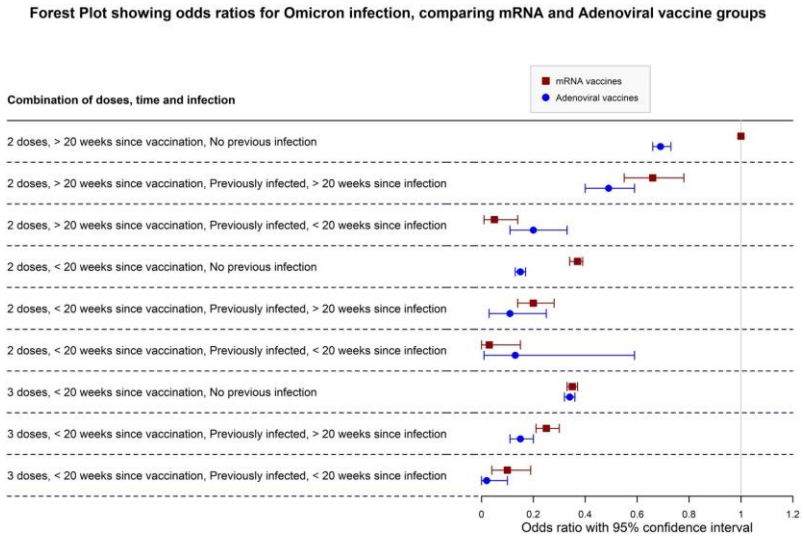
Coronaviruses

Hybrid immunity and protection against infection during the Omicron wave in Malta

John Paul Cauchi, Ausra Dziugyte, Maria-Louise Borg, Tanya Melillo, Graziella Zahra, Christopher Barbara, ...show all

Article: e2156814 | Received 11 Aug 2022, Accepted 05 Dec 2022, Accepted author version posted online: 13 Dec 2022, Published online:02 Jan 2023

Download citation https://doi.org/10.1080/22221751.2022.2156814 Check for updates

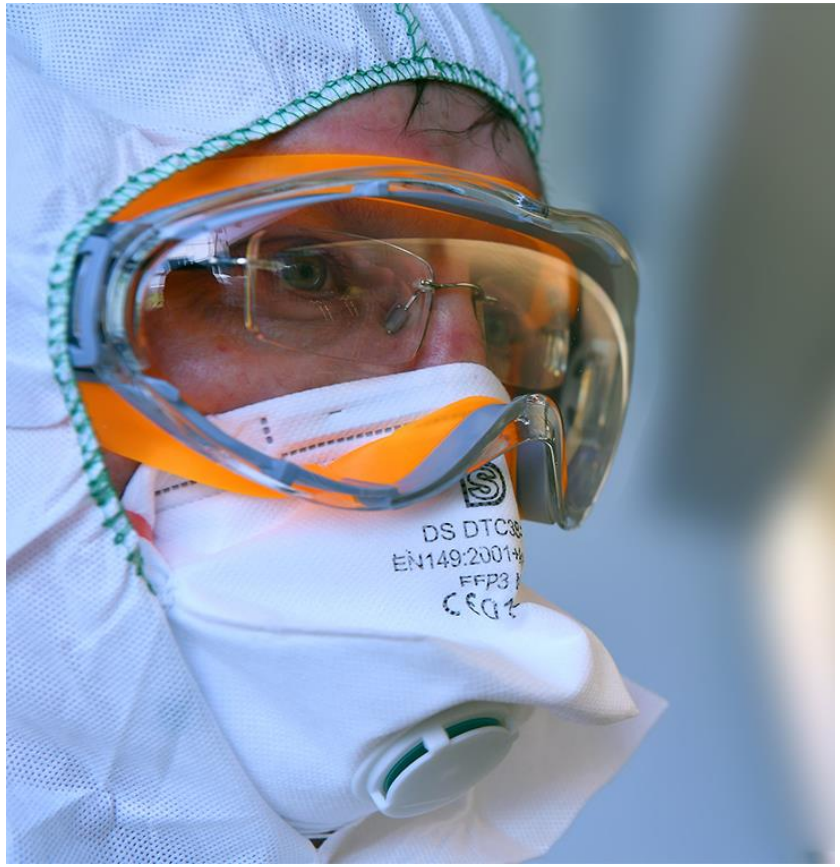


Imaginary Scenario: Pollox, 2025

Imaginary Scenario: Pollox, 2025

Old system

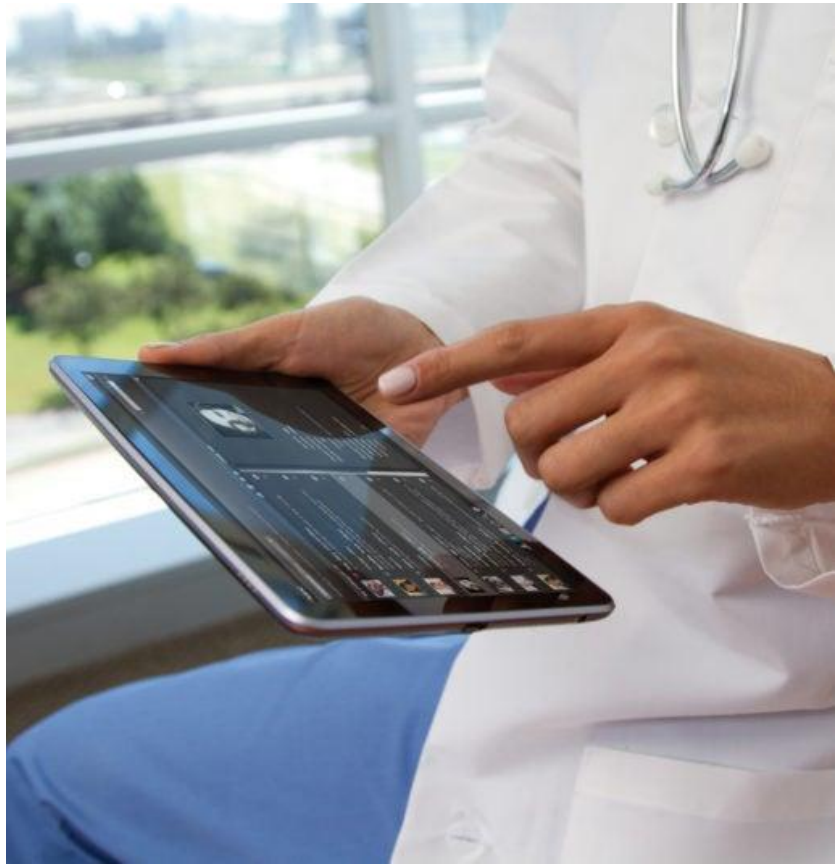
- Data is present in manual files.
- Virtually impossible to ascertain symptom combination to determine presence of disease in people entering hospital
- Have to wait for test kits to arrive from abroad. Considerable impact on economy and mounting anxiety
- Inability to trace the presence of symptom combination
- Mortality increase, draconian actions taken
- Clinicians anxious, unable to be sure who has what



Imaginary Scenario: Pollox, 2025

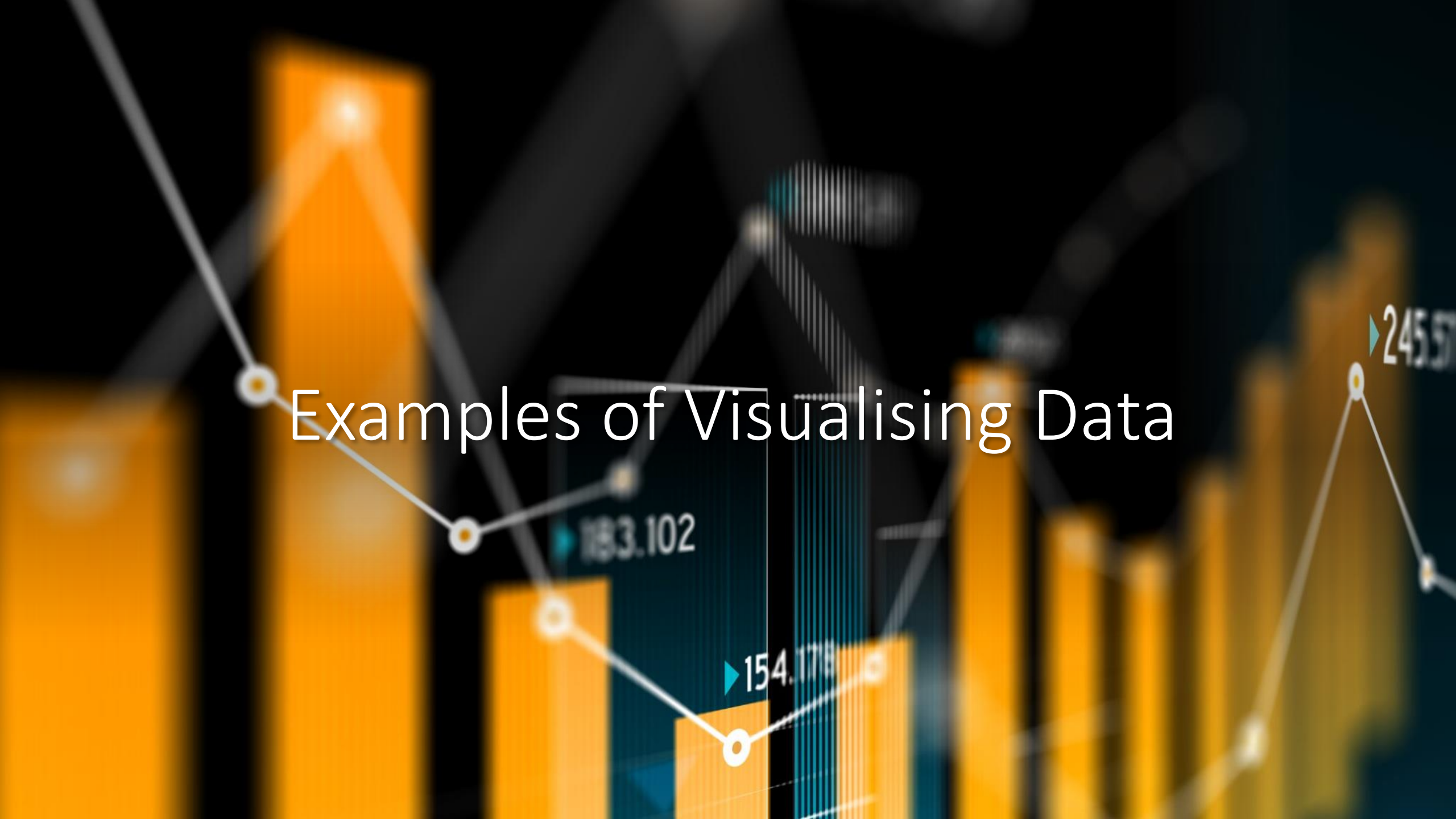
New system

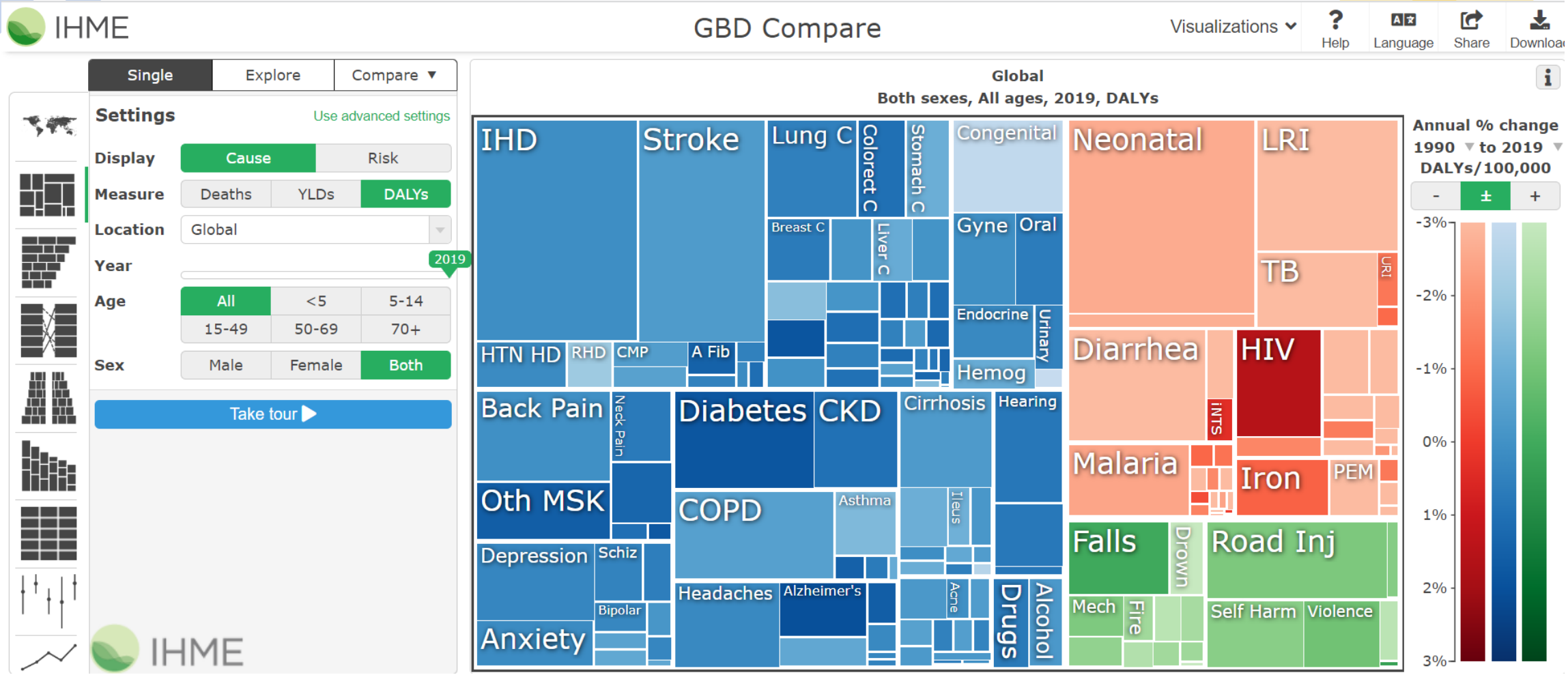
- Data is present in digital files
- Symptom combination is immediately identified, determining presence of disease in people entering hospital
- Measures are taken to immediately quarantine those with symptom combination matching that of Pollox
- Ability to trace symptom combination in clinics across Malta and Gozo
- Prevalence of disease is kept under control
- Clinicians feel reassured, knowing who is likely to be an infected patient



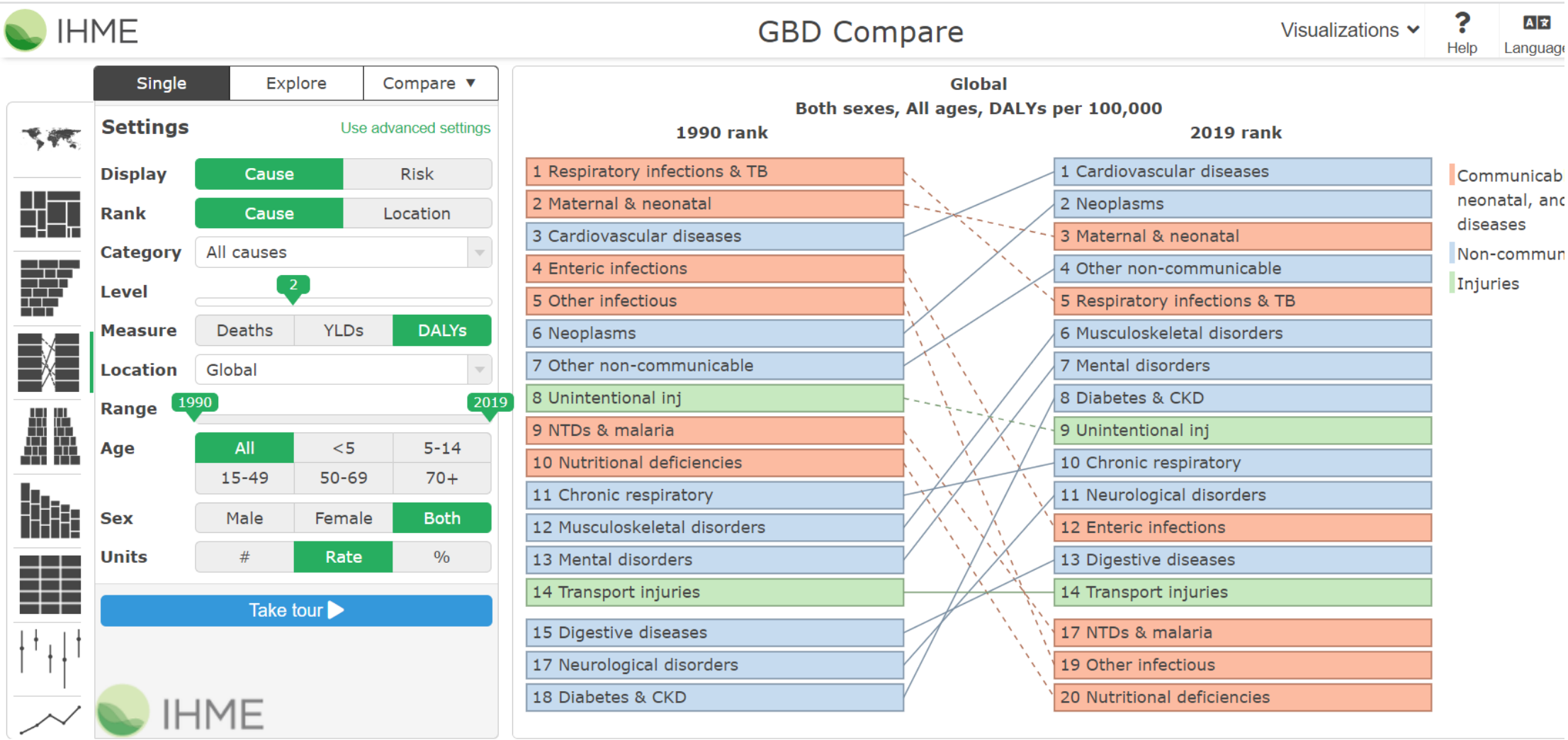
In Summary

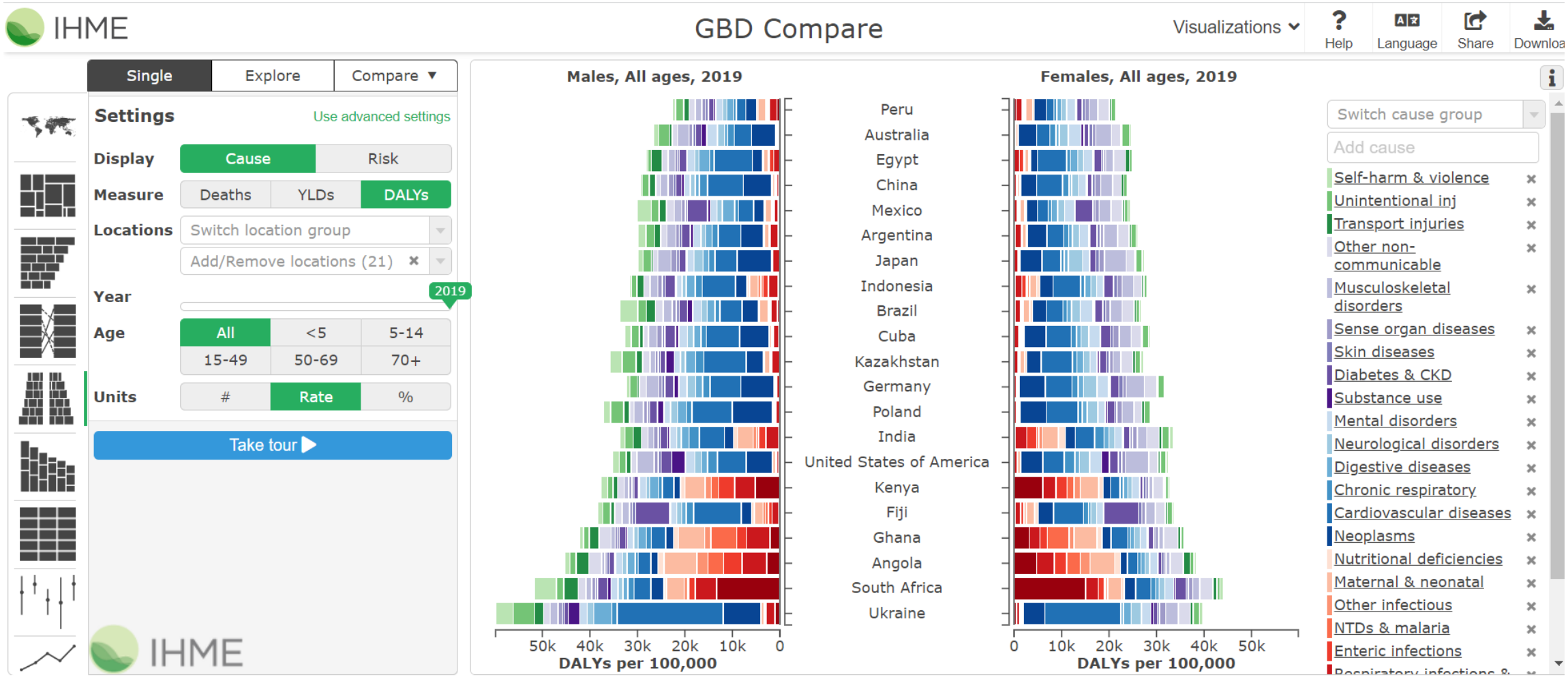
- **Now is an exciting time to enter data visualisation**
- **AI tools, such as Chat GPT,** are immensely powerful and can save a lot of time on searching and finding coding syntax
- Malta's requirements in this sector are **increasing**, especially because there is little to no such capacity in the health sector at this time

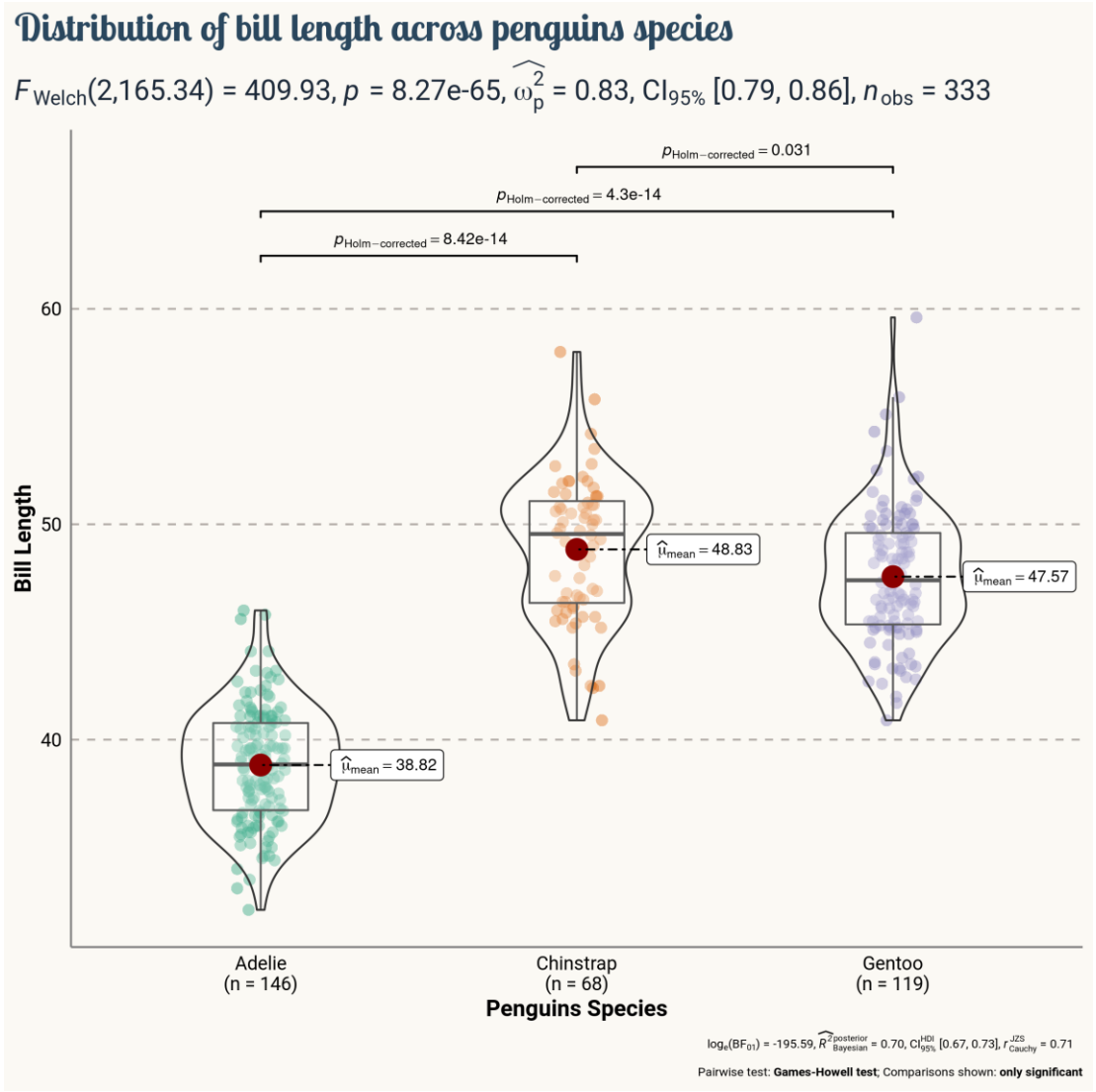

Examples of Visualising Data

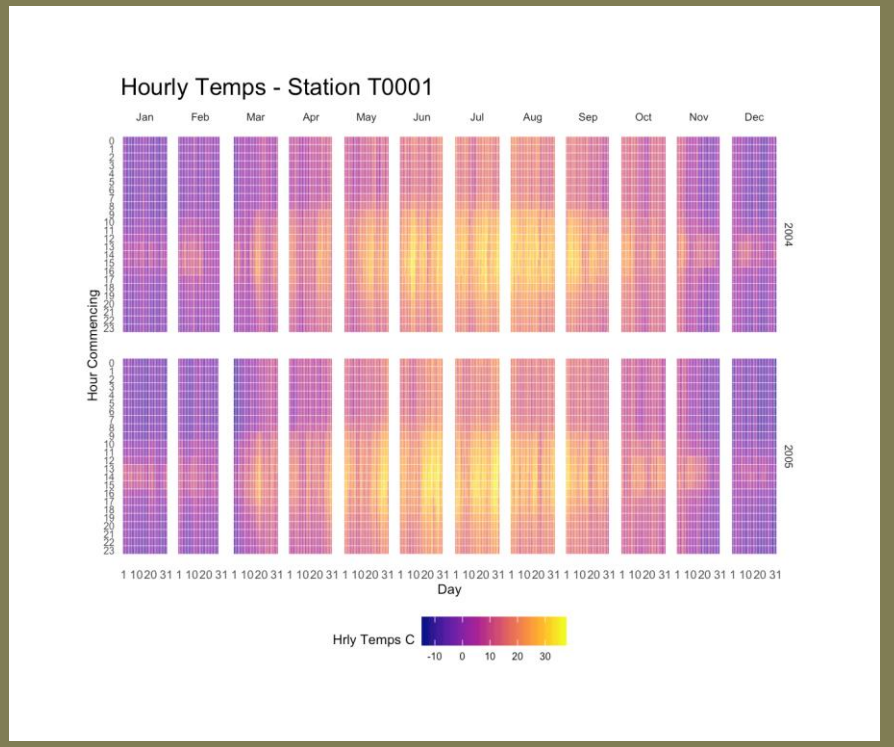


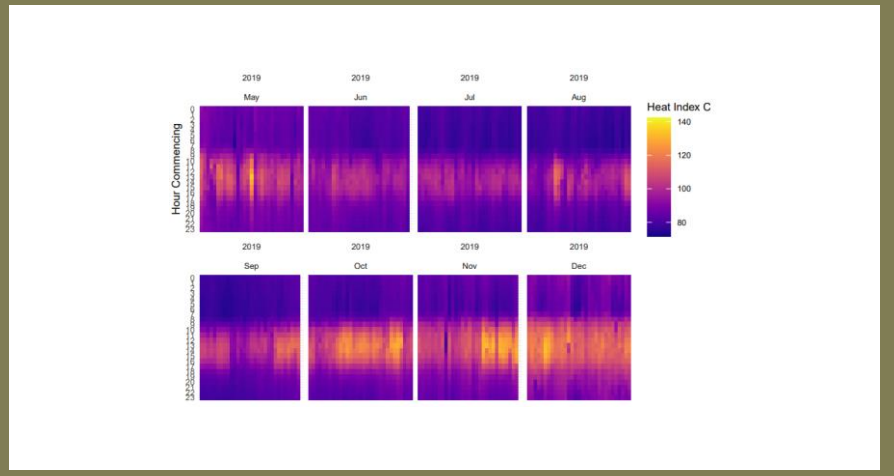
Source: IHME - https://vizhub.healthdata.org/gbd-compare/

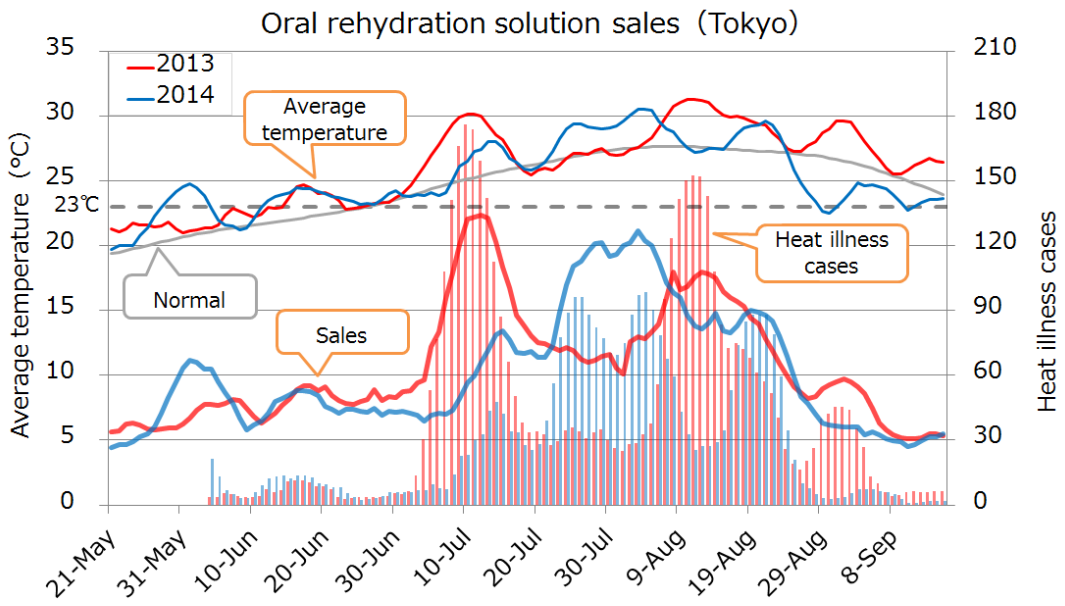

IHME
GBD Compare
Visualizations
Help
Language
Single
Explore
Compare
Settings
Use advanced settings
Display
Cause
Risk
Rank
Cause
Location
Category
All causes
Level
2
Measure
Deaths
YLDs
DALYs
Location
Global
Range
1990
2019
Age
All
<5
5-14
15-49
50-69
70+
Sex
Male
Female
Both
Units
#
Rate
%
Take tour
IHME
Global
Both sexes, All ages, DALYs per 100,000
1990 rank
2019 rank
1 Respiratory infections & TB
2 Maternal & neonatal
3 Cardiovascular diseases
4 Enteric infections
5 Other infectious
6 Neoplasms
7 Other non-communicable
8 Unintentional inj
9 NTDs & malaria
10 Nutritional deficiencies
11 Chronic respiratory
12 Musculoskeletal disorders
13 Mental disorders
14 Transport injuries
15 Digestive diseases
17 Neurological disorders
18 Diabetes & CKD
1 Cardiovascular diseases
2 Neoplasms
3 Maternal & neonatal
4 Other non-communicable
5 Respiratory infections & TB
6 Musculoskeletal disorders
7 Mental disorders
8 Diabetes & CKD
9 Unintentional inj
10 Chronic respiratory
11 Neurological disorders
12 Enteric infections
13 Digestive diseases
14 Transport injuries
17 NTDs & malaria
19 Other infectious
20 Nutritional deficiencies
Communicab
neonatal, anc
diseases
Non-commun
Injuries

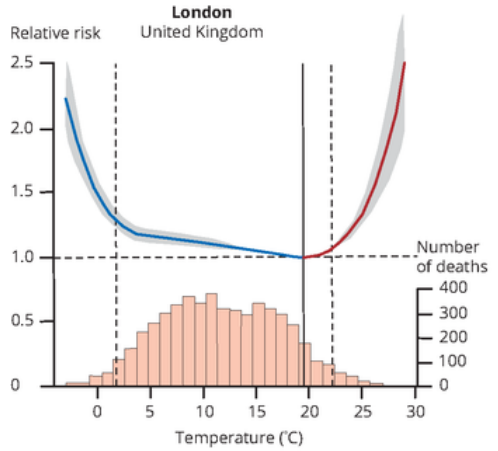

IHME
GBD Compare
Visualizations
Help
Language
Share
Downloa
Single
Explore
Compare
Settings
Use advanced settings
Display
Cause
Risk
Measure
Deaths
YLDs
DALYs
Locations
Switch location group
Add/Remove locations (21)
Year
2019
Age
All
<5
5-14
15-49
50-69
70+
Units
#
Rate
%
Take tour
IHME
Males, All ages, 2019
Females, All ages, 2019
Peru
Australia
Egypt
China
Mexico
Argentina
Japan
Indonesia
Brazil
Cuba
Kazakhstan
Germany
Poland
India
United States of America
Kenya
Fiji
Ghana
Angola
South Africa
Ukraine
50k 40k 30k 20k 10k 0
0 10k 20k 30k 40k 50k
DALYs per 100,000
DALYs per 100,000
Switch cause group
Add cause
Self-harm & violence
Unintentional inj
Transport injuries
Other non-communicable
Musculoskeletal disorders
Sense organ diseases
Skin diseases
Diabetes & CKD
Substance use
Mental disorders
Neurological disorders
Digestive diseases
Chronic respiratory
Cardiovascular diseases
Neoplasms
Nutritional deficiencies
Maternal & neonatal
Other infectious
NTDs & malaria
Enteric infections

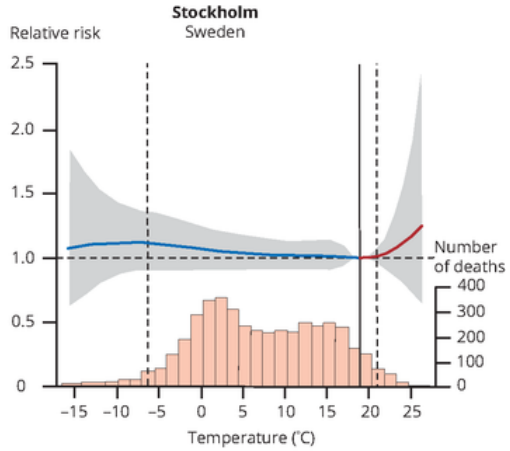

Distribution of bill length across penguins species
$F_{Welch}(2,165.34) = 409.93, p = 8.27e-65, \widehat{\omega}^2_p = 0.83, CI_{95\%}\ [0.79, 0.86], n_{obs} = 333$
$p_{Holm-corrected} = 0.031$
$p_{Holm-corrected} = 4.3e-14$
$p_{Holm-corrected} = 8.42e-14$
Bill Length
60
50
40
$\hat{\mu}_{mean} = 38.82$
$\hat{\mu}_{mean} = 48.83$
$\hat{\mu}_{mean} = 47.57$
Adelie (n = 146)
Chinstrap (n = 68)
Gentoo (n = 119)
Penguins Species
$\log_e(BF_{01}) = -195.59, \widehat{R^2}^{posterior}_{Bayesian} = 0.70, CI^{HDI}_{95\%}\ [0.67, 0.73], r^{JZS}_{Cauchy} = 0.71$
Pairwise test: Games-Howell test; Comparisons shown: only significant

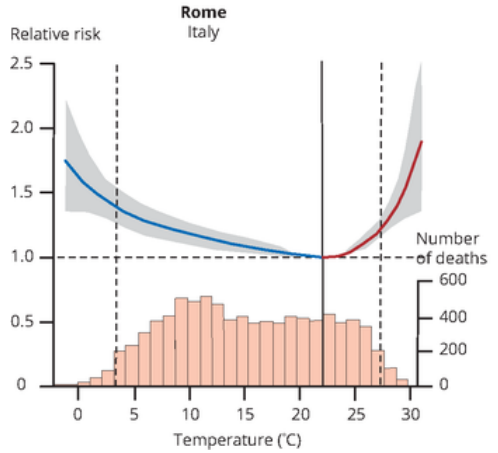

Hourly Temps - Station T0001
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Hour Commencing
2004
2005
1 10 20 31
Day
Hrly Temps C
-10 0 10 20 30

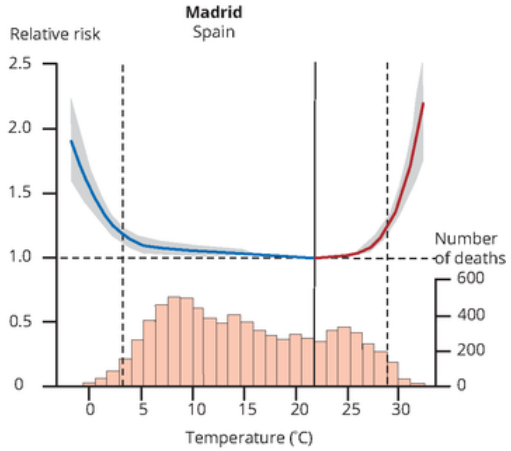

2019
May Jun Jul Aug
Sep Oct Nov Dec
Hour Commencing
Heat Index C
140
120
100
80

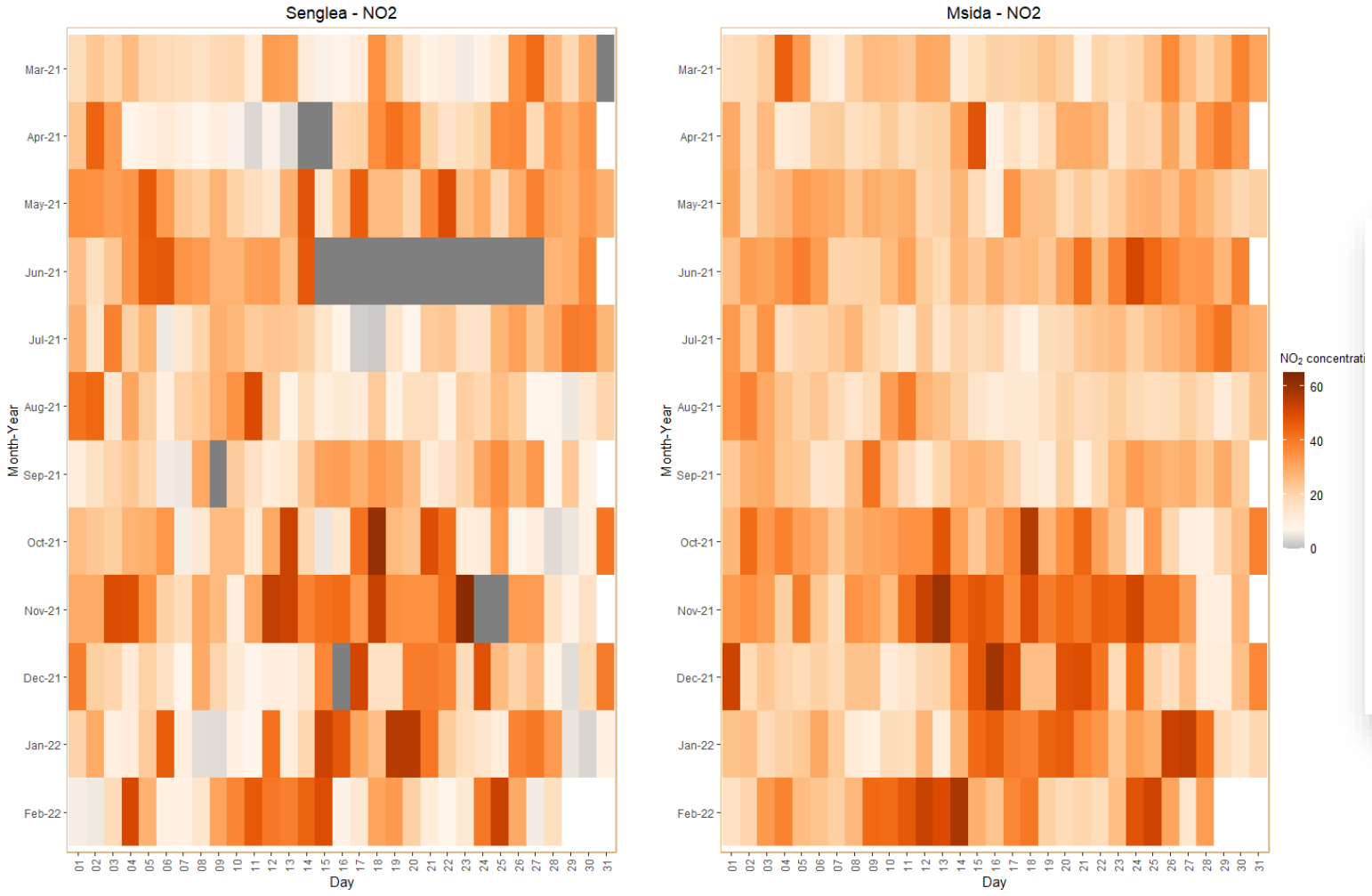

Oral rehydration solution sales (Tokyo)
2013
2014
Average temperature
Normal
Sales
Heat illness cases
23℃
Average temperature (°C)
Heat illness cases
35
30
25
20
15
10
5
0
210
180
150
120
90
60
30
0
21-May
31-May
10-Jun
20-Jun
30-Jun
10-Jul
20-Jul
30-Jul
9-Aug
19-Aug
29-Aug
8-Sep


London
United Kingdom
Relative risk
Number of deaths
Temperature (℃)


Stockholm
Sweden
Relative risk
Number of deaths
Temperature (℃)


Rome
Italy
Relative risk
Number of deaths
Temperature (℃)

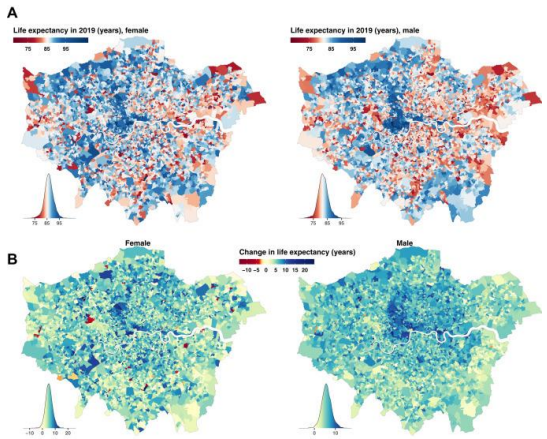

Madrid
Spain
Relative risk
Number of deaths
Temperature (℃)

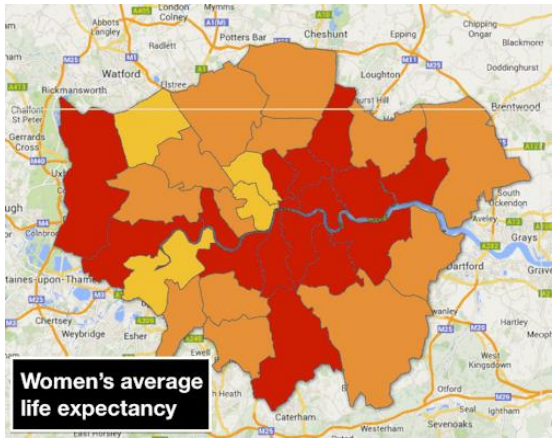


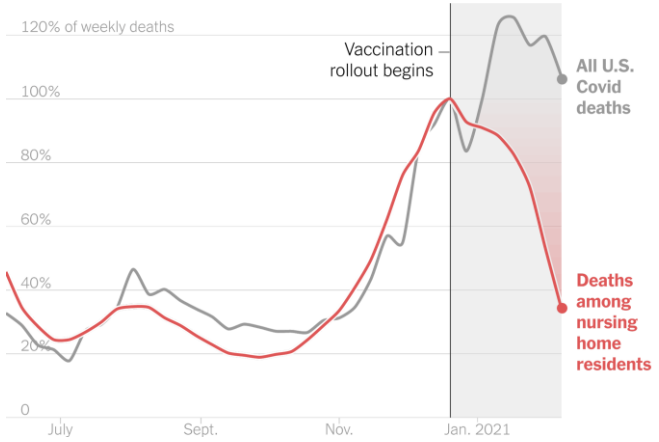
Days when NO2 levels exceeded safe values of 25.0 µg/m³ over a 24-hour period*

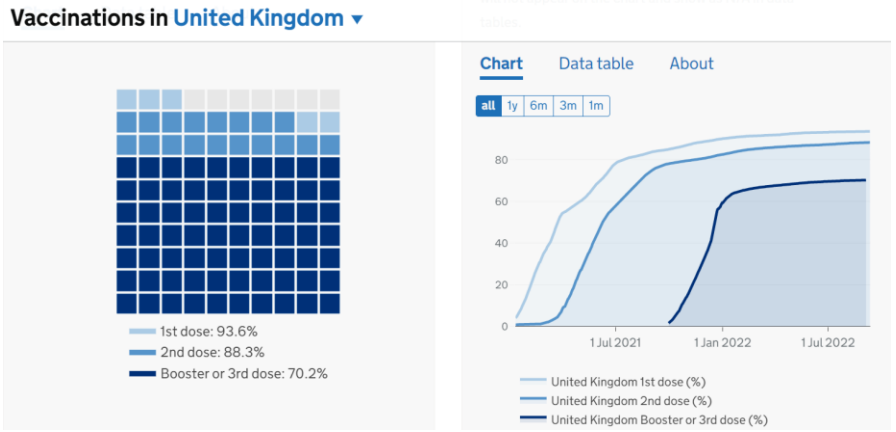
Location	Exceeds limit	Acceptable
Msida	52% (n = 189/365)	48% (n = 176/365)
Senglea	47% (n = 161/346)	53% (n = 185/346)

Levels above 25 µg/m3 (light orange) are dangerous to health…


A
Life expectancy in 2019 (years), female
Life expectancy in 2019 (years), male
B
Female
Change in life expectancy (years)
Male


Women's average life expectancy


120% of weekly deaths
100%
80%
60%
40%
20%
0
Vaccination rollout begins
All U.S. Covid deaths
Deaths among nursing home residents
July
Sept.
Nov.
Jan. 2021


Vaccinations in United Kingdom
1st dose: 93.6%
2nd dose: 88.3%
Booster or 3rd dose: 70.2%
Chart
Data table
About
all
1y
6m
3m
1m
80
60
40
20
0
1 Jul 2021
1 Jan 2022
1 Jul 2022
United Kingdom 1st dose (%)
United Kingdom 2nd dose (%)
United Kingdom Booster or 3rd dose (%)

Conclusion

- Data analysis is key to effective policy decision-making
- Malta needs to invest in its capacity to carry out effective health surveillance
- ‘We’ve always done it this way’ just doesn’t hold water any longer – things are rapidly changing and we need to improve our capabilities
- Centralisation of data makes it easier to carry out effective health data analysis
- Integration with other data sources is possible – geographical, environmental, economic, sociological
- **Result**: better community health care, better health outcomes, cost-saving