OpenSAFELY

NHS England OpenSAFELY GP Data Service

58 million	patients' records
181	projects
31	organisations
>100	completed outputs
42 days	from first shovel to first output
£10 million	spend to date, new funding to scale

GP data presents huge opportunities for the UK

Breadth: one record for every citizen

Depth: detailed info on all GP contacts, diagnoses, tests, more

Research: causes of disease, prevention, treatment benefits / risks / uptake / costs

Improving NHS care: who gets what treatments, where, understand variation.

GP data presents huge challenges

Privacy: records that generate insights also contain everyone's confidential medical history

Transparency: people (professions, campaigners, patients) want to know what's done with their medical records.

Usability: patient records data is hard to work with, and huge, with a historic culture of duplicated effort & closed working.

the data
the outputs
the privacy & transparency tools
the ehr data management tools

Data via OpenSAFELY

Full coded GP records

- Every prescription
- Every blood test and its result
- Every diagnostic code
- Every referral
- More

Additional datasets

- UK Renal Registry
- COVID surveys
- Randomised trial
- Household pseudonym
- More

HES/SUS hospital data

- Every hospital admission
- Every outpatient appointment
- Every A&E attendance

ONS death certificate

and cause of death

		TPP		EMIS		Combined	
Attribute	**Category**	**Total**	**%**	**Total**	**%**	**Total**	**%**
Total		24,011,964	100.0	34,032,530	100.0	58,044,494	100.0
Age group	0-17	4,821,223	20.1	6,901,845	20.3	11,723,068	20.2
	18-24	1,901,509	7.9	2,884,964	8.5	4,786,473	8.2
	25-34	3,340,123	13.9	4,962,526	14.6	8,302,649	14.3
	35-44	3,220,499	13.4	4,745,812	13.9	7,966,311	13.7
	45-54	3,230,861	13.5	4,546,614	13.4	7,777,475	13.4
	55-69	4,202,414	17.5	5,697,231	16.7	9,899,645	17.1
	70-79	2,080,859	8.7	2,699,998	7.9	4,780,857	8.2
	80+	1,214,476	5.1	1,593,540	4.7	2,808,016	4.8
Sex	F	12,004,974	50.0	17,014,169	50.0	29,019,143	50.0
	M	12,006,990	50.0	17,018,361	50.0	29,025,351	50.0
Region	East of England	5,638,753	23.5	1,341,520	3.9	6,980,273	12.0
	East Midlands	4,191,051	17.5	763,830	2.2	4,954,881	8.5
	London	1,702,673	7.1	7,804,070	22.9	9,506,743	16.4
	North East	1,100,356	4.6	1,189,619	3.5	2,289,975	3.9
	North West	2,067,131	8.6	6,875,180	20.2	8,942,311	15.4
	South East	1,582,440	6.6	7,191,261	21.1	8,773,701	15.1
	South West	3,304,393	13.8	2,488,558	7.3	5,792,951	10.0
	West Midlands	988,286	4.1	5,057,090	14.9	6,045,376	10.4
	Yorkshire and The Humber	3,427,713	14.3	1,278,147	3.8	4,705,860	8.1
IMD	1 Most deprived	4,818,642	20.1	7,015,392	20.6	11,834,034	20.4
	2	4,707,307	19.6	7,244,664	21.3	11,951,971	20.6
	3	4,941,725	20.6	6,633,133	19.5	11,574,858	19.9
	4	4,655,595	19.4	6,401,478	18.8	11,057,073	19.0
	5 Least deprived	4,302,292	17.9	6,635,613	19.5	10,937,905	18.8
Ethnicity	White	14,573,038	60.7	17,677,690	51.9	32,250,728	55.6
	Mixed	319,793	1.3	581,965	1.7	901,758	1.6
	South Asian	1,500,012	6.2	2,489,843	7.3	3,989,855	6.9
	Black	515,866	2.1	1,173,341	3.4	1,689,207	2.9
	Other	476,065	2.0	754,993	2.2	1,231,058	2.1

OpenSAFELY >100 outputs

Who gets COVID-19?
Do different ethnic groups have different risks?
Does household transmission play a role?
Do COVID treatments work?
Who gets them?
Do vaccines work?
Who gets them?
Do they have any adverse effects?
What are the consequences of COVID infection?
Who gets “long COVID”?
How did NHS activity change during the pandemic?
How did NHS activity recover after the pandemic?

204 projects across 31 different organisations


OpenSAFELY
Research
opensafely.org/research

nature

Article

Factors associated with COVID-19-related death using OpenSAFELY

https://doi.org/10.1038/s41586-020-2521-4

Received: 15 May 2020

Accepted: 1 July 2020

Published online: 8 July 2020

Check for updates

Elizabeth J. Williamson[1,2], Alex J. Walker[1,2], Krishnan Bhaskaran[1,2], Seb Bacon[1,2], Chris Bates[1,2], Caroline E. Morton[3], Helen J. Curtis[1], Amir Mehrkar[3], David Evans[3], Peter Inglesby[3], Jonathan Cockburn[3], Helen I. McDonald[1,4], Brian MacKenna[3], Laurie Tomlinson[1], Ian J. Douglas[1], Christopher T. Rentsch[1], Rohini Mathur[1], Angel Y. S. Wong[1], Richard Grieve[1], David Harrison[3], Harriet Forbes[1], Anna Schultze[1], Richard Croker[3], John Parry[3], Frank Hester[3], Sam Harper[3], Rafael Perera[3], Stephen J. W. Evans[1], Liam Smeeth[1,4,7] & Ben Goldacre[3,7]✉

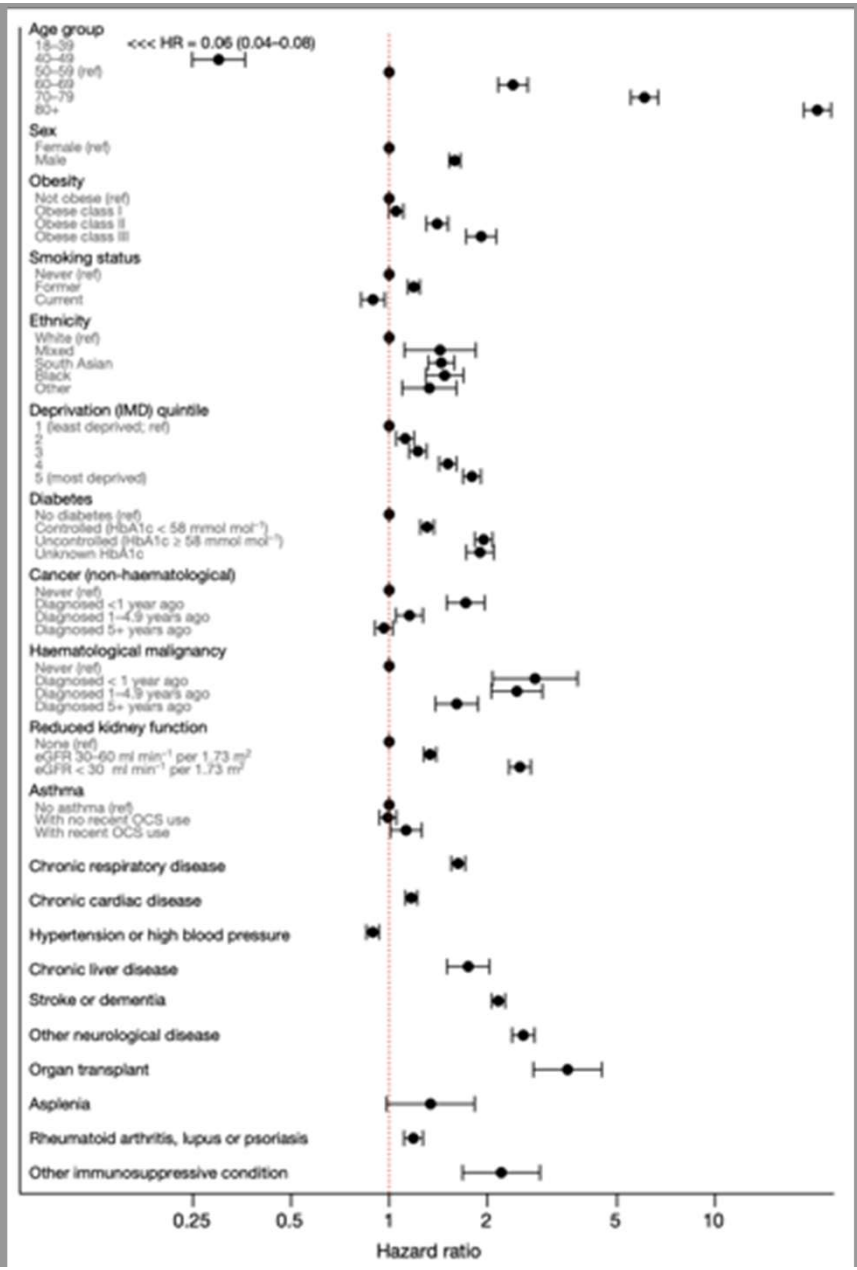
Coronavirus disease 2019 (COVID-19) has rapidly affected mortality worldwide[1]. There is unprecedented urgency to understand who is most at risk of severe outcomes, and this requires new approaches for the timely analysis of large datasets. Working on behalf of NHS England, we created OpenSAFELY—a secure health analytics platform that covers 40% of all patients in England and holds patient data within the existing data centre of a major vendor of primary care electronic health records. Here we used OpenSAFELY to examine factors associated with COVID-19-related death. Primary care records of 17,278,392 adults were pseudonymously linked to 10,926 COVID-19-related deaths. COVID-19-related death was associated with: being male (hazard ratio (HR) 1.59 (95% confidence interval 1.53–1.65)); greater age and deprivation (both with a strong gradient); diabetes; severe asthma; and various other medical conditions. Compared with people of white ethnicity, Black and South Asian people were at higher risk, even after adjustment for other factors (HR 1.48 (1.29–1.69) and 1.45 (1.32–1.58), respectively). We have quantified a range of clinical factors associated with COVID-19-related death in one of the largest cohort studies on this topic so far. More patient records are rapidly being added to OpenSAFELY, we will update and extend our results regularly.

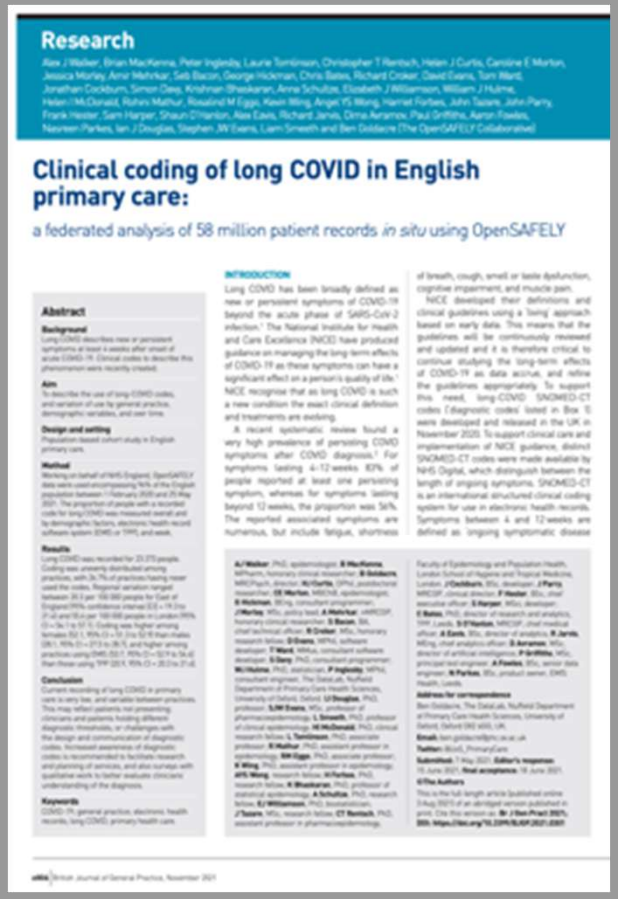
On 11 March 2020, the World Health Organization (WHO) characterized COVID-19—which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)—as a pandemic, after 118,000 cases and 4,291 deaths were reported in 114 countries[2]. As of 6 May 2020 (the date of latest data availability for this study), cases had reached over 3.5 million globally, with more than 240,000 deaths attributed to the virus[3]. On the same day in the UK, there had been 206,715 confirmed cases of COVID-19, and 30,615 COVID-19-related deaths[3].

Age and gender are well-established risk factors for severe COVID-19 outcomes: over 90% of the COVID-19-related deaths in the UK have

in a French intensive-care cohort[7] (n = 124) and a New York hospital presentation cohort[8] (n = 3,615). The risks associated with smoking are unclear[9–11]. People from Black and minority ethnic groups are at increased risk of poor outcomes from COVID-19, for reasons that are unclear[12,13].

Patient care is typically managed through electronic health records, which are commonly used in research. However traditional approaches to the analysis of electronic health records rely on intermittent extracts of small samples of historic data. Evaluating a rapidly arising novel cause of death requires a new approach. We therefore set out to deliver a secure analytics platform inside the data centre of major electronic





Research

Clinical coding of long COVID in English primary care:

a federated analysis of 58 million patient records *in situ* using OpenSAFELY

Abstract

Introduction

Conclusion

Keywords

1st study ever in whole population GP records

RESEARCH

OPEN ACCESS

Association between living with children and outcomes from covid-19: OpenSAFELY cohort study of 12 million adults in England

Abstract

Objective

Design

Setting

Participants

Main outcome measures

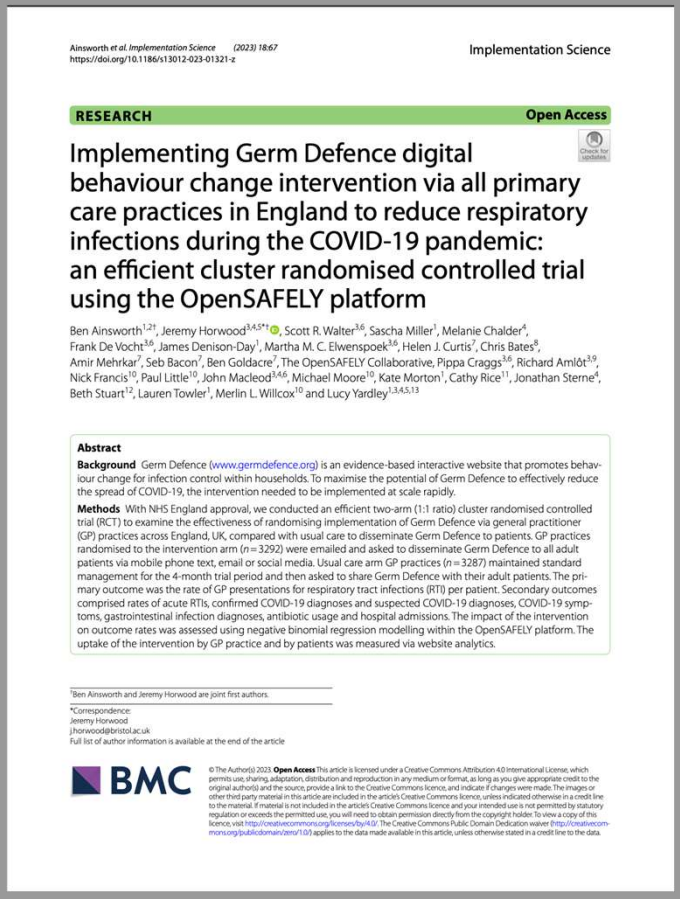
Results

Conclusions

What is already known on this topic

What this study adds

Household pseudo-identifier for household risks



Implementation Science

RESEARCH Open Access

Implementing Germ Defence digital behaviour change intervention via all primary care practices in England to reduce respiratory infections during the COVID-19 pandemic: an efficient cluster randomised controlled trial using the OpenSAFELY platform

Abstract

Background

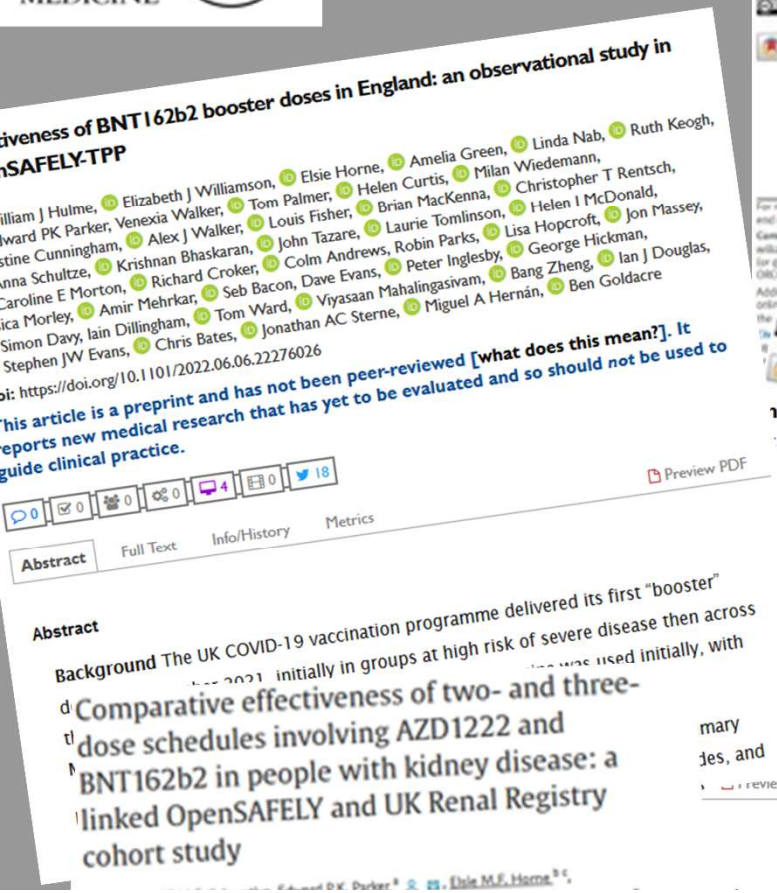
Methods

24 million patient randomised trial, follow up in TRE

Do vaccines wane over time?
How helpful are boosters?

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

University of BRISTOL



iveness of BNT162b2 booster doses in England: an observational study in
nSAFELY-TPP

illiam J Hulme, Elizabeth J Williamson, Elsie Horne, Amelia Green, Linda Nab, Ruth Keogh,
dward PK Parker, Venexia Walker, Tom Palmer, Helen Curtis, Milan Wiedemann,
stine Cunningham, Alex J Walker, Louis Fisher, Brian MacKenna, Christopher T Rentsch,
nna Schultze, Krishnan Bhaskaran, John Tazare, Laurie Tomlinson, Helen I McDonald,
aroline E Morton, Richard Croker, Colm Andrews, Robin Parks, Lisa Hopcroft, Jon Massey,
ica Morley, Amir Mehrkar, Seb Bacon, Dave Evans, Peter Inglesby, George Hickman,
Simon Davy, Iain Dillingham, Tom Ward, Viyasaan Mahalingasivam, Bang Zheng, Ian J Douglas,
Stephen JW Evans, Chris Bates, Jonathan AC Sterne, Miguel A Hernán, Ben Goldacre

oi: https://doi.org/10.1101/2022.06.06.22276026

his article is a preprint and has not been peer-reviewed [what does this mean?]. It
eports new medical research that has yet to be evaluated and so should not be used to
guide clinical practice.

0 0 0 4 0 18

Abstract Full Text Info/History Metrics Preview PDF

Abstract

Background The UK COVID-19 vaccination programme delivered its first "booster"
initially in groups at high risk of severe disease then across
was used initially, with

Comparative effectiveness of two- and three-dose schedules involving AZD1222 and BNT162b2 in people with kidney disease: a linked OpenSAFELY and UK Renal Registry cohort study

mary
des, and
Preview PDF

OPEN ACCESS

Check for updates

Comparative effectiveness of BNT162b2 versus mRNA-1273 covid-19 vaccine boosting in England: matched cohort study in OpenSAFELY-TPP

William J Hulme,[1] Elsie M F Horne,[2,3] Edward P K Parker,[4] Ruth H Keogh,[4] Elizabeth J Williamson,[4] Venexia Walker,[2,5] Tom M Palmer,[2,5] Helen J Curtis,[1] Alex J Walker,[1] Colm D Andrews,[1] Amir Mehrkar,[1] Jessica Morley,[1] Brian MacKenna,[1] Sebastian C J Bacon,[1] Ben Goldacre,[1] Miguel A Hernán,[4,6] Jonathan A C Sterne[2,3,6]

ABSTRACT

OBJECTIVE

To compare the effectiveness of the BNT162b2 mRNA (Pfizer-BioNTech) and mRNA-1273 (Moderna) covid-19 vaccines during the booster programme in England.

DESIGN

Matched cohort study, emulating a comparative effectiveness trial.

interval 163.3 to 165.1) for BNT162b2 and 159.9 (159.0 to 160.8) for mRNA-1273; the hazard ratio comparing mRNA-1273 with BNT162b2 was 0.95 (95% confidence interval 0.95 to 0.96). The 20 week risks per 1000 for hospital admission with covid-19 were 0.75 (0.71 to 0.79) for BNT162b2 and 0.65 (0.61 to 0.69) for mRNA-1273; the hazard ratio was 0.89 (0.82 to 0.95). Covid-19 related deaths were rare: the 20 week risks per 1000 were 0.028 (0.021 to 0.037) for BNT162b2 and 0.024 (0.018 to 0.033) for mRNA-1273; hazard ratio 0.83 (0.58 to 1.19). Comparative effectiveness was generally similar within

OPEN ACCESS

Check for updates

Waning effectiveness of BNT162b2 and ChAdOx1 covid-19 vaccines over six months since second dose: OpenSAFELY cohort study using linked electronic health records

Elsie M F Horne,[1,2] William J Hulme,[1] Ruth H Keogh,[4] Tom M Palmer,[1,5] Elizabeth J Williamson,[4] Edward P K Parker,[4] Amelia Green,[1] Venexia Walker,[1,5] Alex J Walker,[1] Helen Curtis,[1] Louis Fisher,[1] Brian MacKenna,[1] Amir Mehrkar,[1] Sebastian Bacon,[1] Richard Croker,[1] Lisa Hopcroft,[1] Robin Y Park,[1] Jon Massey,[1] Jessica Morley,[1] George Hickman,[1] Simon Davy,[1] David Evans,[1] Peter Inglesby,[1] Caroline E Morton,[1] Tom Ward,[1] Iain Dillingham,[1] Ben Goldacre,[1] Miguel A Hernán,[4,6] Jonathan A C Sterne[1,3,6]

ABSTRACT

OBJECTIVE

To estimate waning of covid-19 vaccine effectiveness over six months after second dose.

DESIGN

Cohort study, approved by NHS England.

SETTING

Linked primary care, hospital, and covid-19 records within the OpenSAFELY-TPP database.

PARTICIPANTS

Adults without previous SARS-CoV-2 infection were eligible, excluding care home residents and healthcare professionals.

EXPOSURES

effectiveness was quantified as ratios of adjusted hazard ratios per four w
subgroups aged ≥65 ye
vulnerable, 40-64 year

RESULTS

1 951 866 and 3 219 3
two doses of BNT162b
and 2 622 980 remain
vaccine effectiveness
across outcomes and
years subgroup, ratios
for covid-19 related h
related death, and po
from 1.19 (95

Annals of Internal Medicine

Challenges in Estimating the Effectiveness of Using Observational Data

William J. Hulme, PhD; Elizabeth J. Williamson, PhD; Alex J. Walker, PhD; Amir Mehrkar, PhD; [illegible]

RESEARCH AND REPORTING METHODS

Comparative effectiveness of ChAdOx1 versus BNT162b2 covid-19 vaccines in health and social care workers in England: cohort study using OpenSAFELY

William J Hulme,[1] Elizabeth J Williamson,[2] Amelia C A Green,[1] Krishnan Bhaskaran,[2] Helen I McDonald,[2] Christopher T Rentsch,[2] Anna Schultze,[2] John Tazare,[2] Helen J Curtis,[1] Alex J Walker,[1] Laurie A Tomlinson,[2] Tom Palmer,[3,4] Elsie M F Horne,[4,5] Brian MacKenna,[1] Caroline E Morton,[1] Amir Mehrkar,[1] Jessica Morley,[1] Louis Fisher,[1] Sebastian C J Bacon,[1] David Evans,[1] Peter Inglesby,[1] George Hickman,[1] Simon Davy,[1] Tom Ward,[1] Richard Croker,[1] Rosalind M Eggo,[2] Angel YS Wong,[2] Rohini Mathur,[2] Kevin Wing,[2] Harriet Forbes,[2] Daniel J Grint,[2] Ian J Douglas,[2] Stephen JW Evans,[2] Liam Smeeth,[2] Chris Bates,[4] Jonathan Cockburn,[4] John Parry,[4] Frank Hester,[4] Sam Harper,[4] Jonathan AC Sterne,[4,5,6] Miguel A Hernán,[4,6] Ben Goldacre[1]

ABSTRACT

OBJECTIVE

To compare the effectiveness of the BNT162b2 mRNA (Pfizer-BioNTech) and the ChAdOx1 (Oxford-AstraZeneca) covid-19 vaccines against infection and covid-19 disease in health and social care workers.

DESIGN

Cohort study, emulating a comparative effectiveness trial, on behalf of NHS England.

SETTING

Linked primary care, hospital, and covid-19 surveillance records available within the OpenSAFELY-TPP research platform, covering a period when the SARS-CoV-2 Alpha variant was dominant.

PARTICIPANTS

317 341 health and socialcare workers vaccinated between 4 January and 28 February 2021, registered with a general practice using the TPP SystmOne clinical information system in England, and not clinically extremely vulnerable.

INTERVENTIONS

Vaccination with either BNT162b2 or ChAdOx1 administered as part of the national covid-19 vaccine roll-out.

MAIN OUTCOME MEASURES

Recorded SARS-CoV-2 positive test, or covid-19 related attendance at an accident and emergency (A&E) department or hospital admission occuring within 20 weeks of receipt of the first vaccine dose.

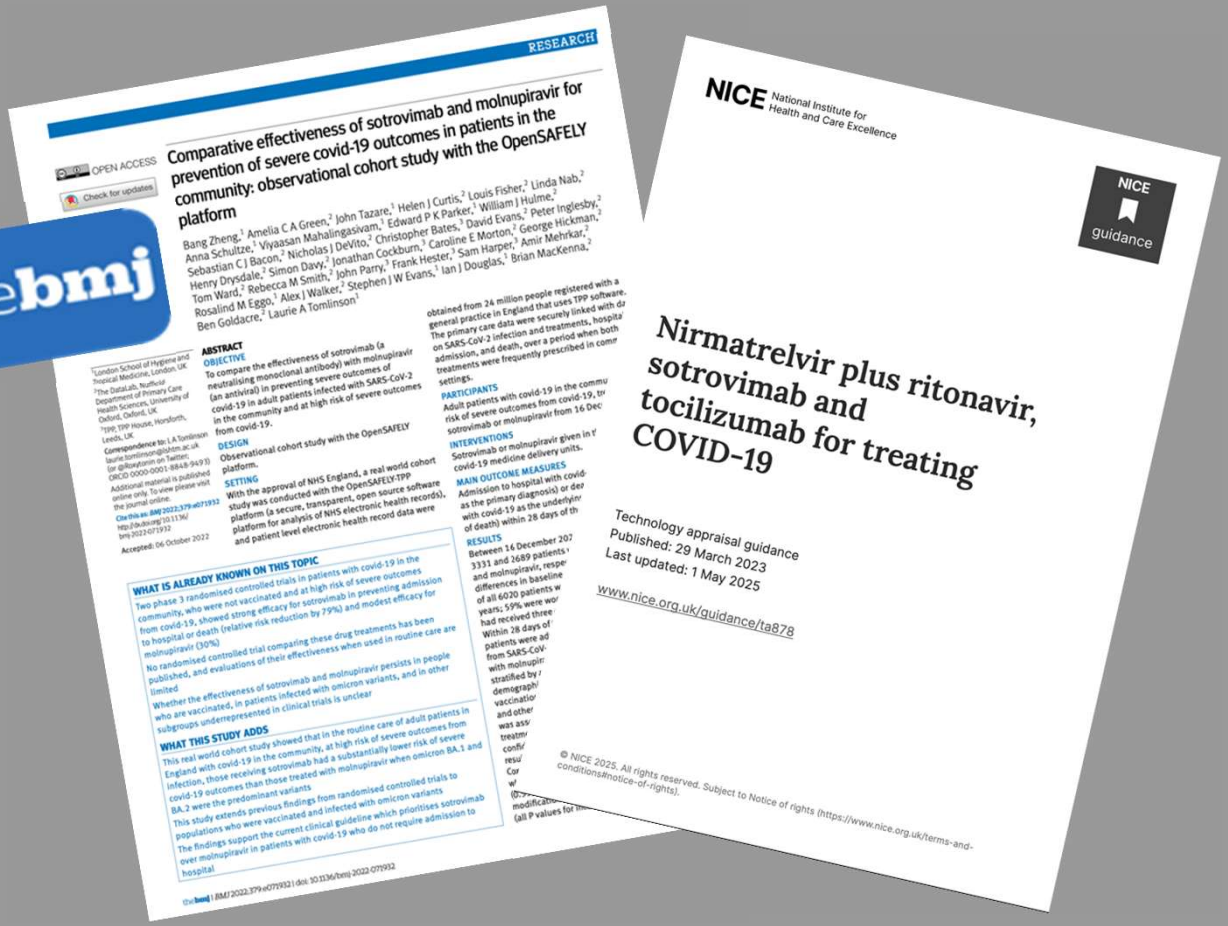
RESULTS

Over the duration of 118 771 person-years of follow-up there were 6962 positive SARS-CoV-2 tests, 282 covid-19 related A&E attendances, and 166 covid-19 related hospital admissions. The cumulative incidence of each outcome was similar for both vaccines during the first 20 weeks after vaccination. The cumulative incidence of recorded SARS-CoV-2 infection 20-weeks after first dose vaccination with BNT162b2 was 21.7 per 1000 people (95% confidence interval 20.9 to 22.4) and with ChAdOx1 was 23.7 (21.8 to 25.6), representing a difference of 2.06 per 1000 people (0.04 to 4.04). The difference in the cumulative incidence per 1000 people of covid-19 related A&E attendance at 20 weeks was 0.06 per 1000 people (95% CI −0.31 to 0.43). For covid-19 related hospital admission, this difference was 0.11 per 1000 people (−0.22 to 0.44).

CONCLUSIONS

In this cohort of healthcare workers where we would not anticipate vaccine type to be related to health status, we found no substantial differences in the
SARS-CoV-2 infection or covid-19 disease
dropped

Do the new covid treatments work?



RESEARCH

OPEN ACCESS

Comparative effectiveness of sotrovimab and molnupiravir for prevention of severe covid-19 outcomes in patients in the community: observational cohort study with the OpenSAFELY platform

Bang Zheng,[1] Amelia C A Green,[2] John Tazare,[1] Helen J Curtis,[2] Louis Fisher,[1] Linda Nab,[2] Anna Schultze,[1] Vivaasan Mahalingasivam,[1] Edward P K Parker,[1] William J Hulme,[2] Sebastian C J Bacon,[2] Nicholas J DeVito,[2] Christopher Bates,[3] David Evans,[2] Peter Inglesby,[2] Henry Drysdale,[2] Simon Davy,[2] Jonathan Cockburn,[3] Caroline E Morton,[2] George Hickman,[2] Tom Ward,[2] Rebecca M Smith,[2] John Parry,[3] Frank Hester,[3] Sam Harper,[3] Amir Mehrkar,[2] Rosalind M Eggo,[1] Alex J Walker,[2] Stephen J W Evans,[1] Ian J Douglas,[2] Brian MacKenna,[2] Ben Goldacre,[2] Laurie A Tomlinson[1]

ABSTRACT

OBJECTIVE
To compare the effectiveness of sotrovimab (a neutralising monoclonal antibody) with molnupiravir (an antiviral) in preventing severe outcomes of covid-19 in adult patients infected with SARS-CoV-2 in the community and at high risk of severe outcomes from covid-19.

DESIGN
Observational cohort study with the OpenSAFELY platform.

SETTING
With the approval of NHS England, a real world cohort study was conducted with the OpenSAFELY-TPP platform (a secure, transparent, open source software platform for analysis of NHS electronic health records), and patient level electronic health record data were

obtained from 24 million people registered with a general practice in England that uses TPP software. The primary care data were securely linked with da on SARS-CoV-2 infection and treatments, hospita admission, and death, over a period when both treatments were frequently prescribed in comr settings.

PARTICIPANTS
Adult patients with covid-19 in the commu risk of severe outcomes from covid-19, tr sotrovimab or molnupiravir from 16 Dec

INTERVENTIONS
Sotrovimab or molnupiravir given in r covid-19 medicine delivery units.

MAIN OUTCOME MEASURES
Admission to hospital with covid- as the primary diagnosis) or dea with covid-19 as the underlyin of death) within 28 days of th

RESULTS
Between 16 December 207 3331 and 2689 patients and molnupiravir, respe differences in baseline of all 6030 patients w years; 59% were wo had received three Within 28 days of patients were ad from SARS-CoV- with molnupir stratified by demographi vaccinatio and other was ass treatm confi resu Cor w (95 (all P values for

WHAT IS ALREADY KNOWN ON THIS TOPIC

Two phase 3 randomised controlled trials in patients with covid-19 in the community, who were not vaccinated and at high risk of severe outcomes from covid-19, showed strong efficacy for sotrovimab in preventing admission to hospital or death (relative risk reduction by 79%) and modest efficacy for molnupiravir (30%)

No randomised controlled trial compaling these drug treatments has been published, and evaluations of their effectiveness when used in routine care are limited

Whether the effectiveness of sotrovimab and molnupiravir persists in people who are vaccinated, in patients infected with omicron variants, and in other subgroups underrepresented in clinical trials is unclear

WHAT THIS STUDY ADDS

This real world cohort study showed that in the routine care of adult patients in England with covid-19 in the community, at high risk of severe outcomes from infection, those receiving sotrovimab had a substantially lower risk of severe covid-19 outcomes than those treated with molnupiravir when omicron BA.1 and BA.2 were the predominant variants

This study extends previous findings from randomised controlled trials to populations who were vaccinated and infected with omicron variants

The findings support the current clinical guideline which prioritises sotrovimab over molnupiravir in patients with covid-19 who do not require admission to hospital

thebmj | BMJ 2022;379:e071932 | doi: 10.1136/bmj-2022-071932

NICE National Institute for Health and Care Excellence

NICE guidance

Nirmatrelvir plus ritonavir, sotrovimab and tocilizumab for treating COVID-19

Technology appraisal guidance
Published: 29 March 2023
Last updated: 1 May 2025

www.nice.org.uk/guidance/ta878

Who is getting them?

OPEN ACCESS

bmjmedicine

ORIGINAL RESEARCH

Trends, variation, and clinical characteristics of recipients of antiviral drugs and neutralising monoclonal antibodies for covid-19 in community settings: retrospective, descriptive cohort study of 23.4 million people in OpenSAFELY

Amelia C A Green,[1] Helen J Curtis,[1] Rose Higgins,[1] Linda Nab,[1] Vivaasan Mahalingasivam,[2] Rebecca M Smith,[1] Amir Mehrkar,[1] Peter Inglesby,[1] Henry Drysdale,[1] Nicholas J DeVito,[1] Richard Croker,[1] Christopher T Rentsch,[2] Krishnan Bhaskaran,[2] John Tazare,[2] Bang Zheng,[2] Colm D Andrews,[1] Sebastian C J Bacon,[1] Simon Davy,[1] Iain Dillingham,[1] David Evans,[1] Louis Fisher,[1] George Hickman,[1] Lisa E M Hopcroft,[1] William J Hulme,[1] Jon Massey,[1] Orla MacDonald,[1] Jessica Morley,[1] Caroline E Morton,[1] Robin Y Park,[1] Alex J Walker,[1] Tom Ward,[1] Milan Wiedemann,[1] Christopher Bates,[3] Jonathan Cockburn,[3] John Parry,[3] Frank Hester,[3] Sam Harper,[3] Ian J Douglas,[2] Stephen J W Evans,[2] Ben Goldacre,[2] Laurie A Tomlinson,[2] Brian MacKenna[2]

ABSTRACT

OBJECTIVE To ascertain patient eligibility status and describe coverage of antiviral drugs and neutralising monoclonal antibodies (nMAB) as treatment for covid-19 in community settings in England.

DESIGN Retrospective, descriptive cohort study, approved by NHS England.

SETTING Routine clinical data from 23.4 million people linked to data on covid-19 infection and treatment, within the OpenSAFELY-TPP database.

PARTICIPANTS Outpatients with covid-19 at high risk of severe outcomes.

INTERVENTIONS Nirmatrelvir/ritonavir (paxlovid), sotrovimab, molnupiravir, casirivimab/imdevimab, or remdesivir, used in the community by covid-19 medicine delivery units.

RESULTS 91 870 outpatients with covid-19 were identified between 21 December 2021 and 28 April 2022 to be at high risk of severe outcomes and therefore potentially eligible for antiviral or nMAB treatment (or both). Of these patients, 19 040 (20%) received treatment (sotrovimab, 9660 (51%); molnupiravir, 4620 (24%); paxlovid, 4680 (25%); casirivimab/imdevimab, 50 (<1%); and remdesivir, 30 (<1%)). The proportion of patients treated increased from 9% (190/2220) in the first week of treatment availability to 29% (460/1600) in the latest week. The proportion treated varied by high risk group, being lowest in those with liver disease (16%; 95% confidence interval 15% to 17%); by treatment type, with sotrovimab favoured over molnupiravir and paxlovid in all but three high risk groups (Down's syndrome (35%; 30% to 39%), rare neurological conditions (43%; 43% to 47%), and immune deficiencies (48%; 47% to 50%)); by age, ranging from 480 years (53%; 52% to 54%) to 50-59 years (23%; 22% to 23%); by ethnic group, ranging from black (11%; 10% to 12%) to white (21%; 21% to 21%); by NHS region, ranging from 13% (12% to 14%) in Yorkshire and the Humber to 25% (24% to 25%) in the East of England); and by deprivation level, ranging from 15% (14% to 15%) in the most deprived areas to 23% (23% to 24%) in the least deprived areas. Groups that also had lower coverage included unvaccinated patients (7%; 6% to 9%), those with dementia (6%; 5% to 7%), and care home residents (6%; 6% to 7%).

CONCLUSIONS Using the OpenSAFELY platform, we were able to identify patients with covid-19 at high risk of severe outcomes who were potentially eligible to receive treatment and assess the coverage of these new treatments among these patients. In the context of a rapid deployment of a new service, the NHS analytical code used to determine eligibility could have been over-inclusive

WHAT IS ALREADY KNOWN ABOUT THIS TOPIC

⇒ Since the emergence of covid-19, several approaches to treatment have been tried and evaluated

⇒ These approaches have mainly consisted of treatments such as dexamethasone, which were used in UK hospitals, from early on in the pandemic to prevent progression to severe disease

⇒ Until recently (December 2021), no treatments have been widely used in community settings across England

WHAT THIS STUDY ADDS

⇒ After the roll-out of antiviral drugs and neutralising monoclonal antibodies as treatment for patients with covid-19, patients who were potentially eligible to receive such treatments were identified and the coverage of these new treatments was assessed among these patients, in as close to real time as the available data flows would support

⇒ While the proportion of the potentially eligible patients receiving treatment increased over time, rising from 9% (190/2220) in the first week of the roll-out to 29% (460/1600) in the last week of April 2022, coverage varied between key clinical, geographical, and population subgroups

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY

⇒ Targeted activity might be needed to resolve lower treatment rates observed among certain geographical areas and key groups including ethnic group, people living in areas of higher deprivation, and care home residents

BMJ

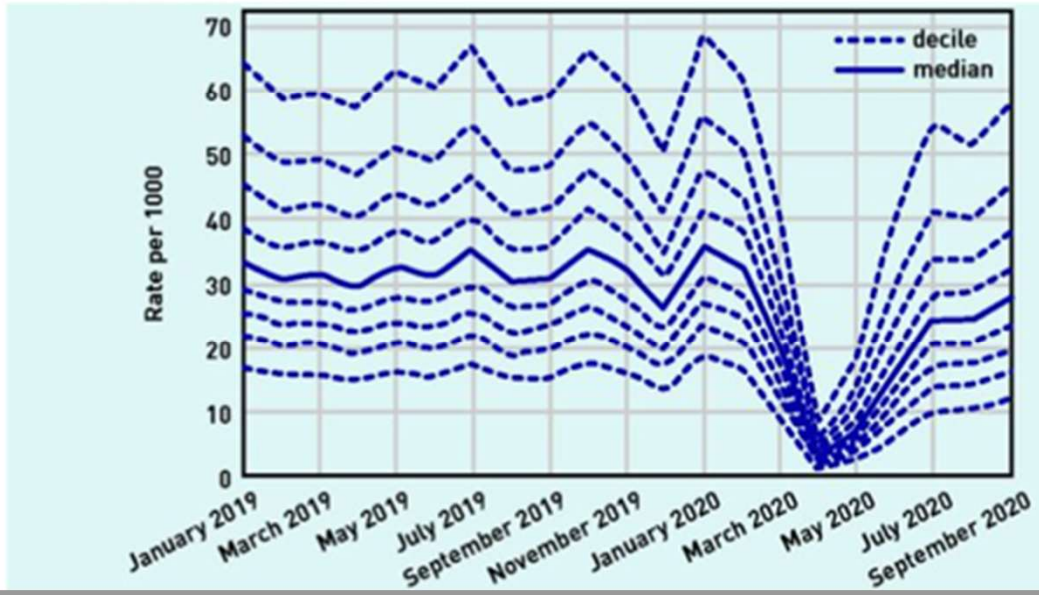
Green ACA, et al. BMJMED-2023-2. doi:10.1136/bmjmed-2022-000276

1

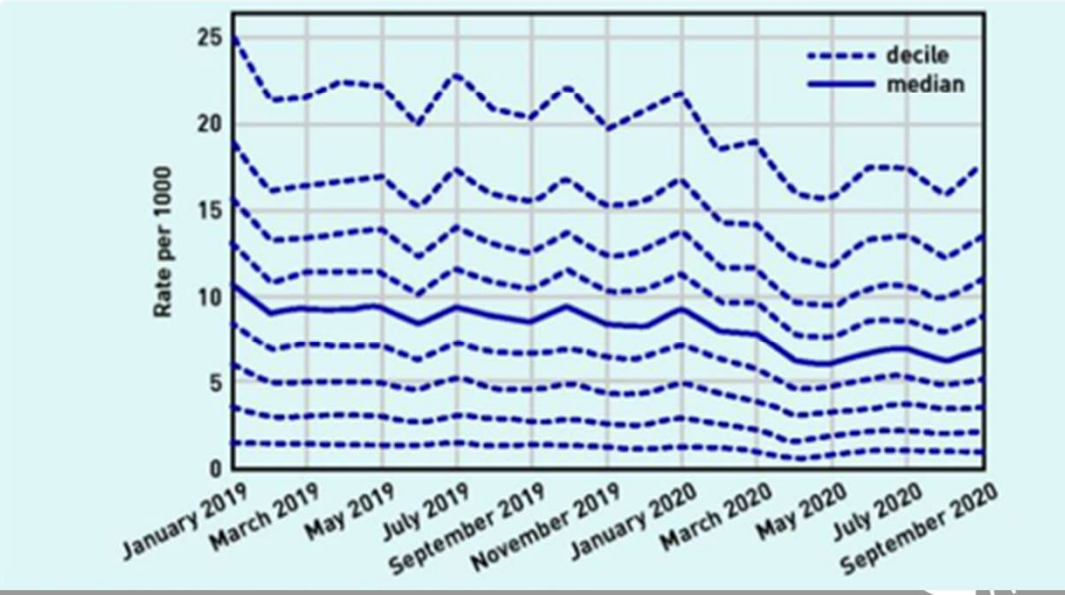

THE Sun
SCHOOLS 'NO RISK' Schools do not fuel
BBC NEWS
Coronavirus: Living with
The New York Time
Study of 17 Mil
The Telegraph
Hydroxychloroquine does not pre
exposure to Co
INDEPENDENT
Coronavirus: Underlying health
conditions don't explain higher BAM
Mail Online
Parents of schoolchildren
a greater risk of being

Monitoring NHS activity and outcomes in near real time

Serum cholesterol tests performed per 1000 members of the population



Blood coagulation tests performed per 1000 members of the population



OpenSAFELY

Home

Contact

REPORTS

COVID Treatments

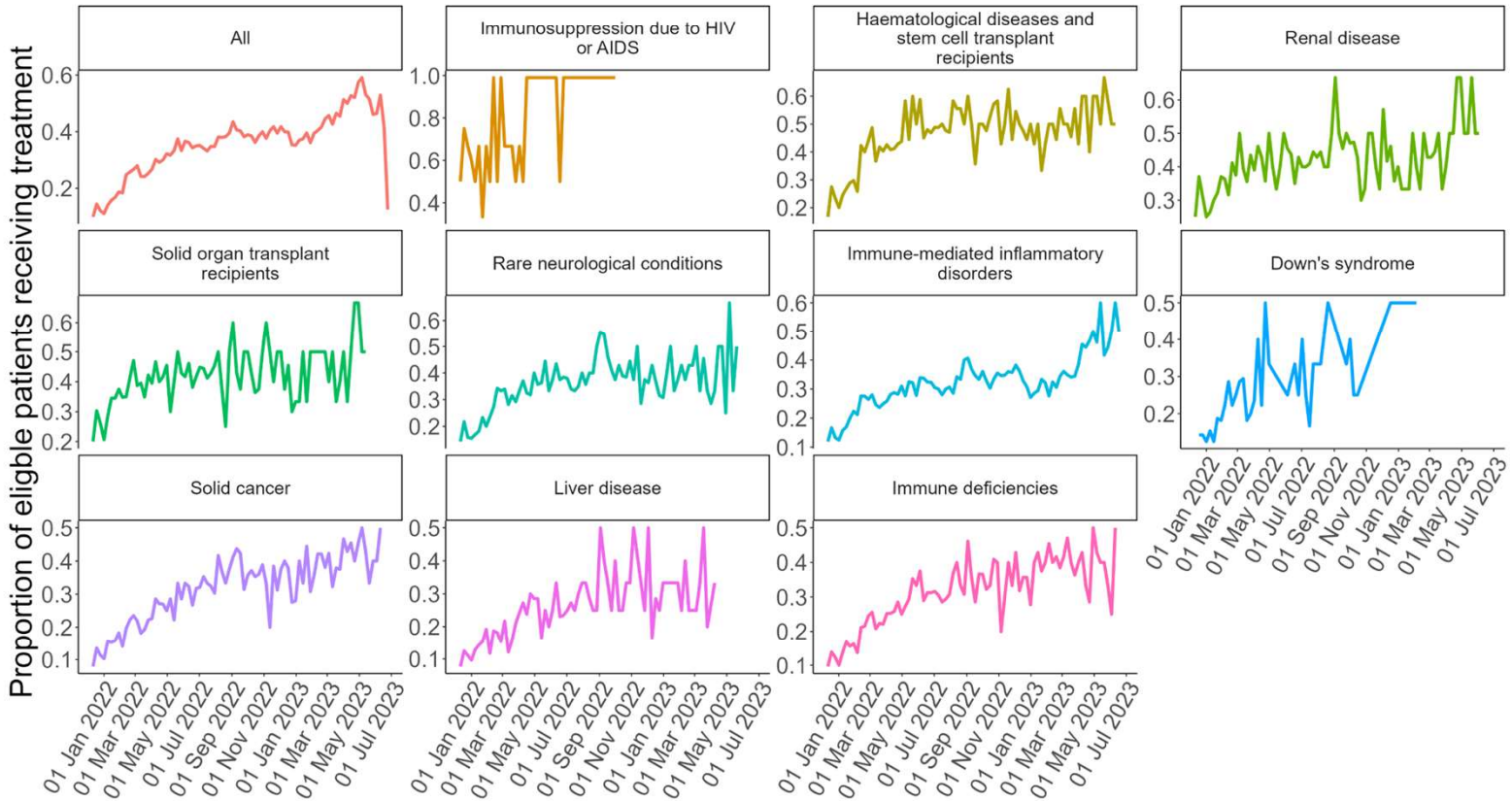
Community Antiviral/MAB coverage report

Databases

Reports

Vaccine Reports

Figure 3 Weekly proportion of eligible patients receiving an antiviral or nMAB for treating COVID-19 since 11th December 2021, stratified by high risk cohort. Proportions are rounded to 2 decimal places.



Key demographic and clinical characteristics of treated patients

the data
the outputs
the privacy & transparency tools
the ehr data management tools

Trust

NHSE and GP as data controller
Records don’t move
Users work on dummy data
Total, “code-level” transparency

Trust

NHSE and GP as data controller

Records don’t move

Users work on dummy data

Total, code-level transparency

Trust

NHSE and GP as data controller
Records don’t move
Users work on dummy data
Total, code-level transparency

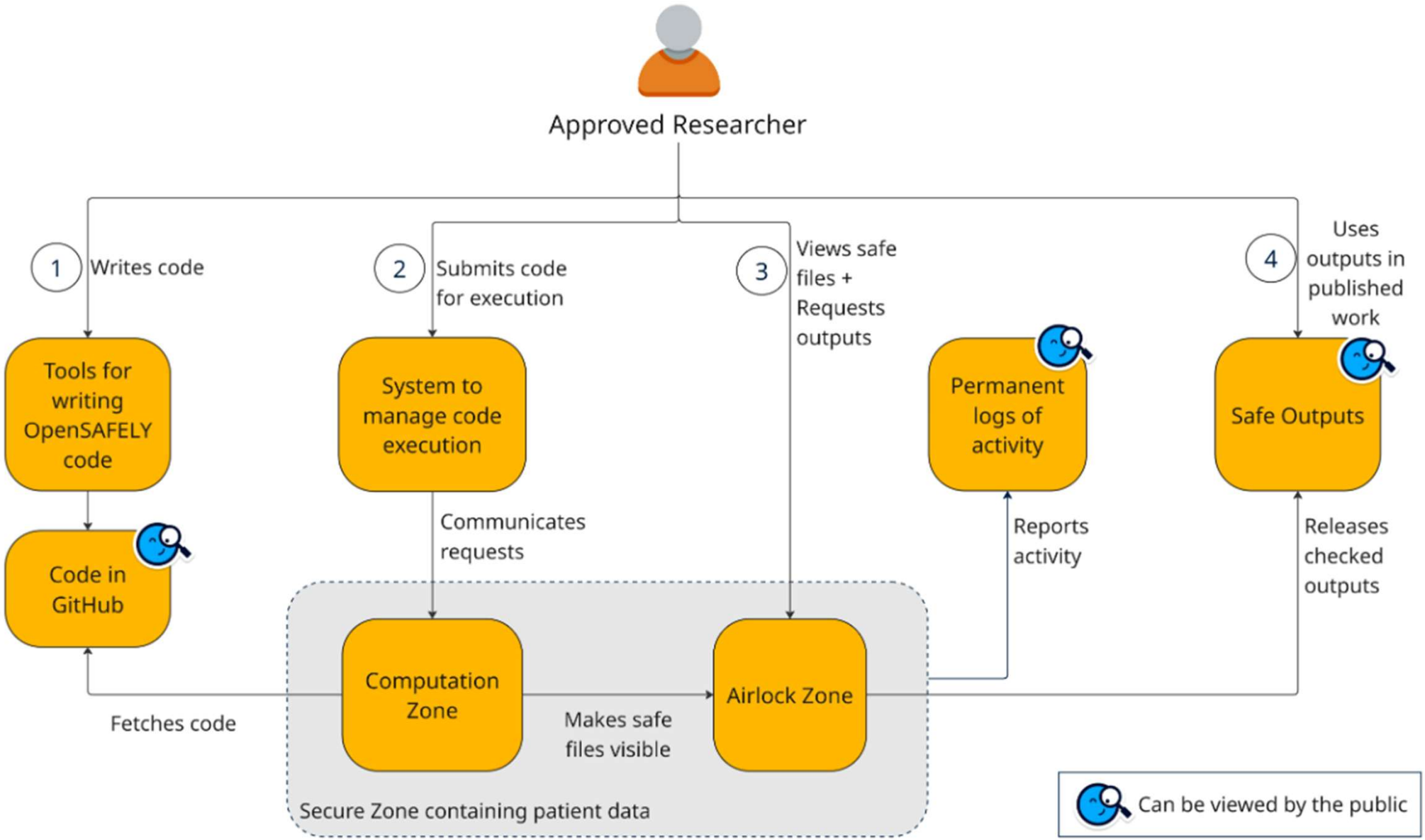
Trust

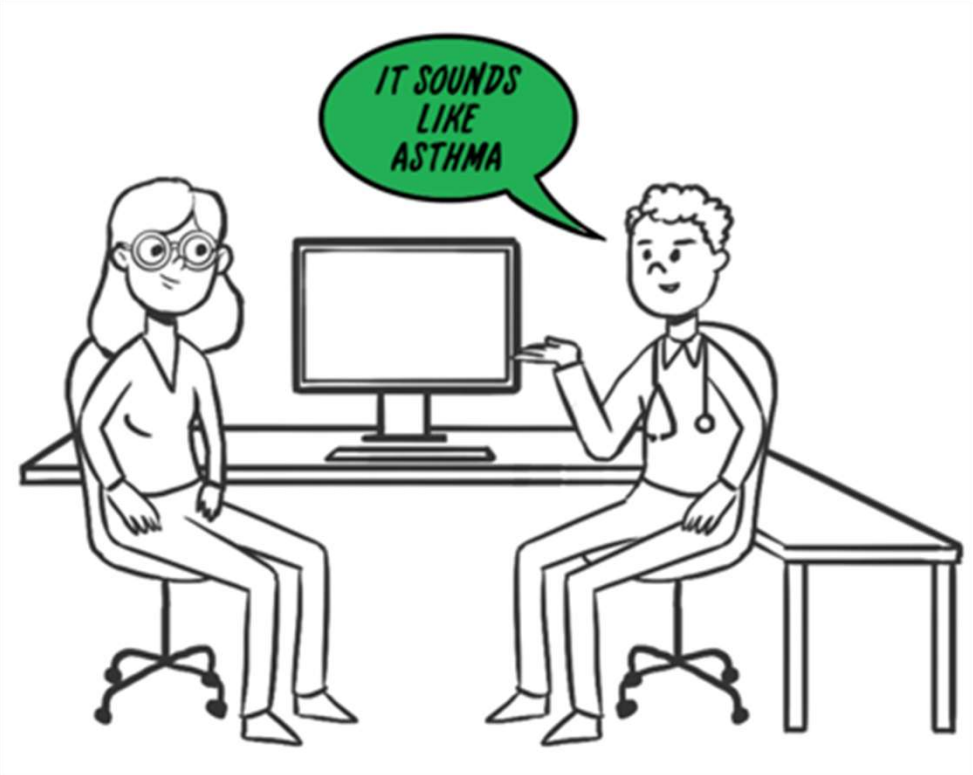
NHSE and GP as data controller

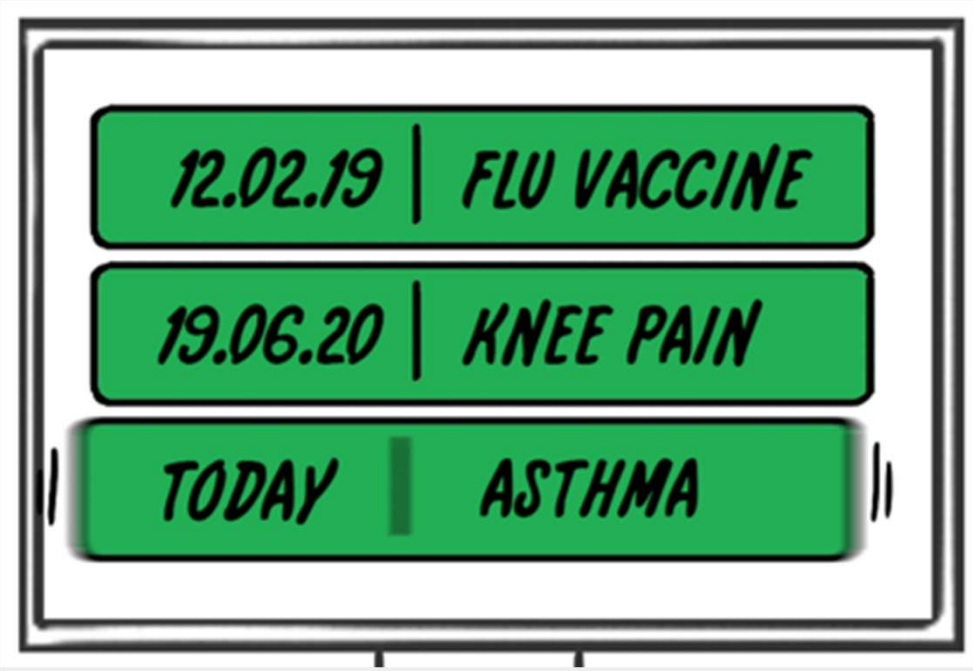
Records don’t move

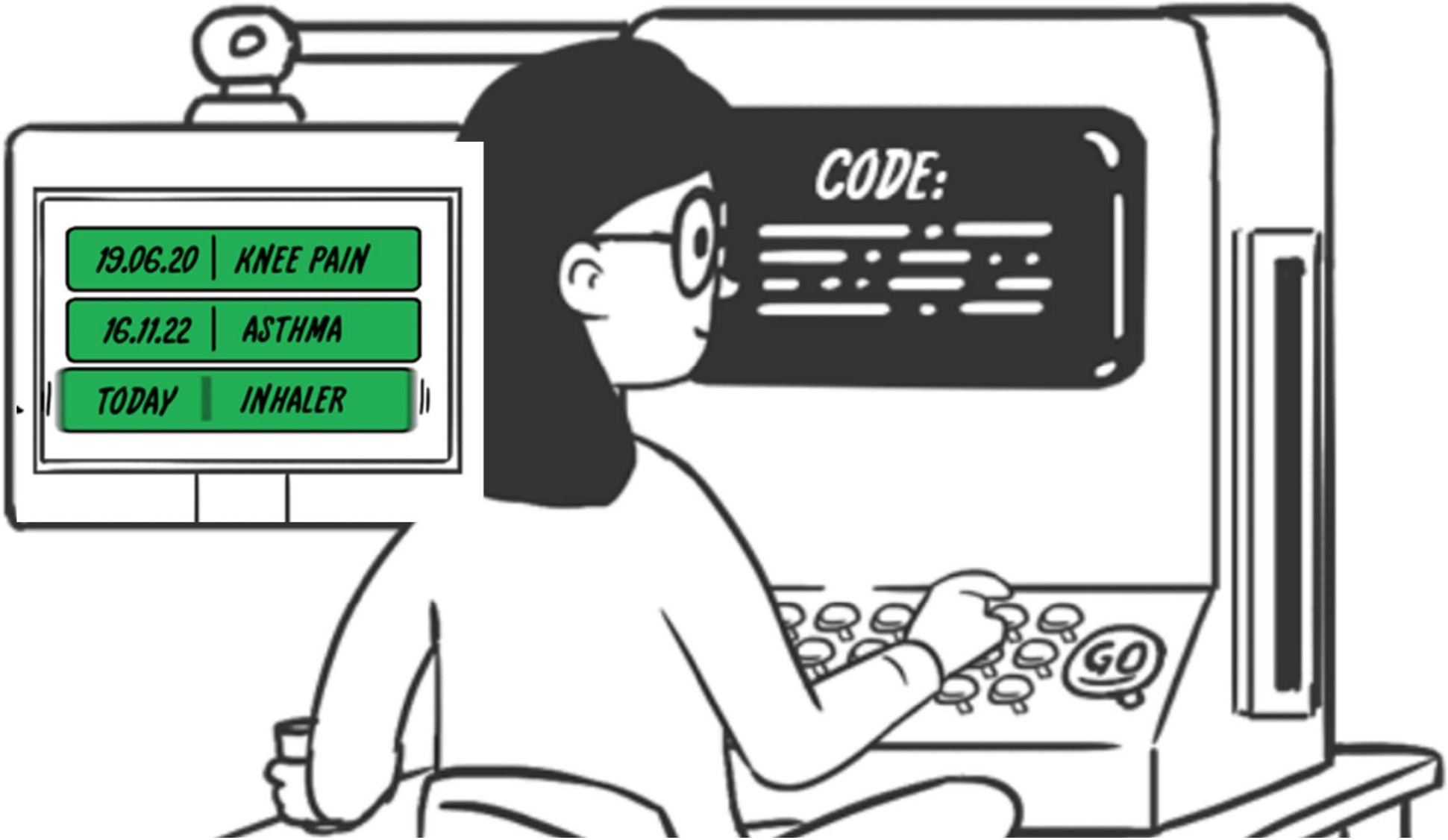
Users work on dummy data

Total, code-level transparency

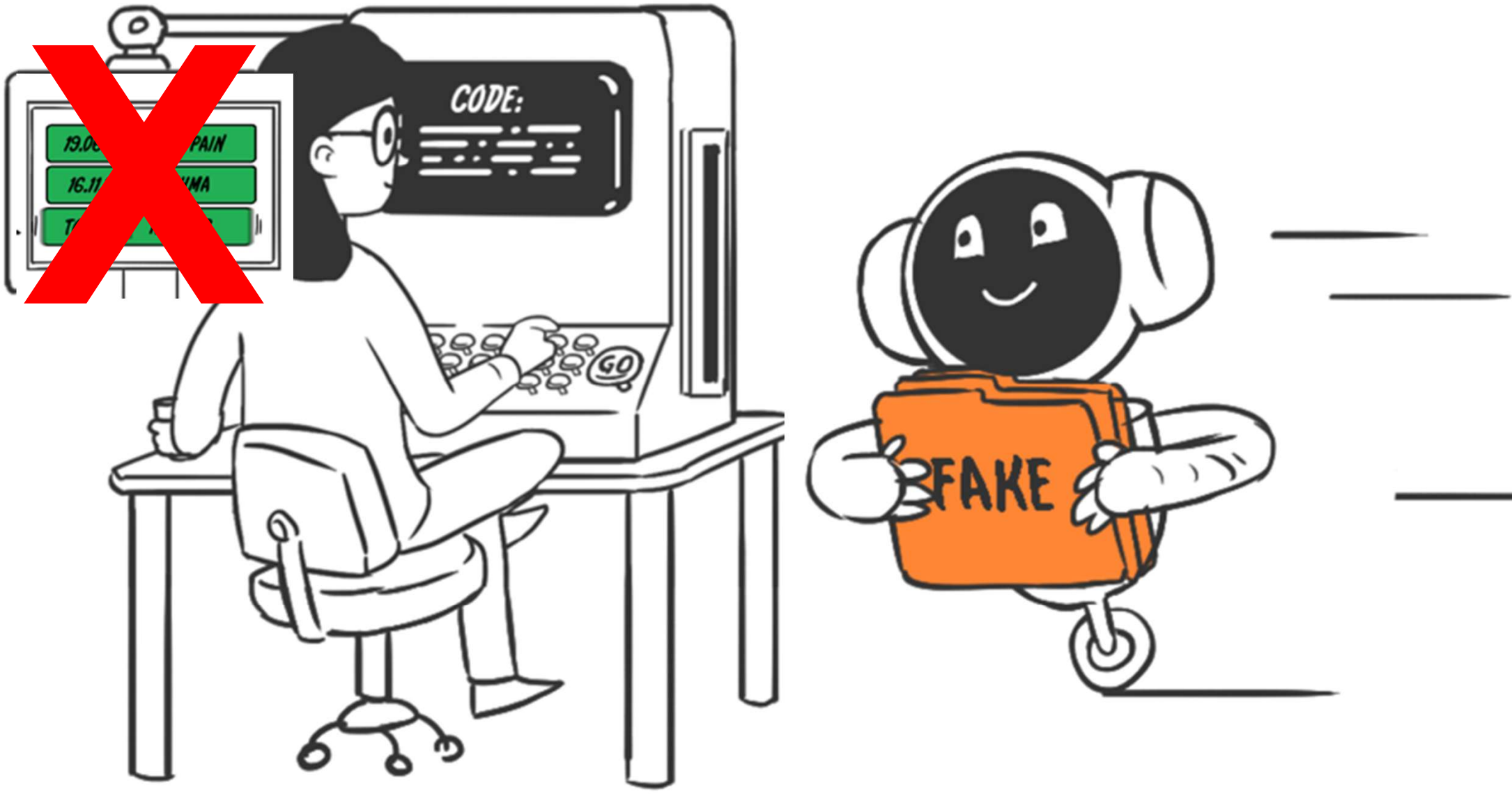

Approved Researcher
1
Writes code
2
Submits code for execution
3
Views safe files + Requests outputs
4
Uses outputs in published work
Tools for writing OpenSAFELY code
System to manage code execution
Permanent logs of activity
Safe Outputs
Code in GitHub
Communicates requests
Reports activity
Releases checked outputs
Computation Zone
Airlock Zone
Fetches code
Makes safe files visible
Secure Zone containing patient data
Can be viewed by the public

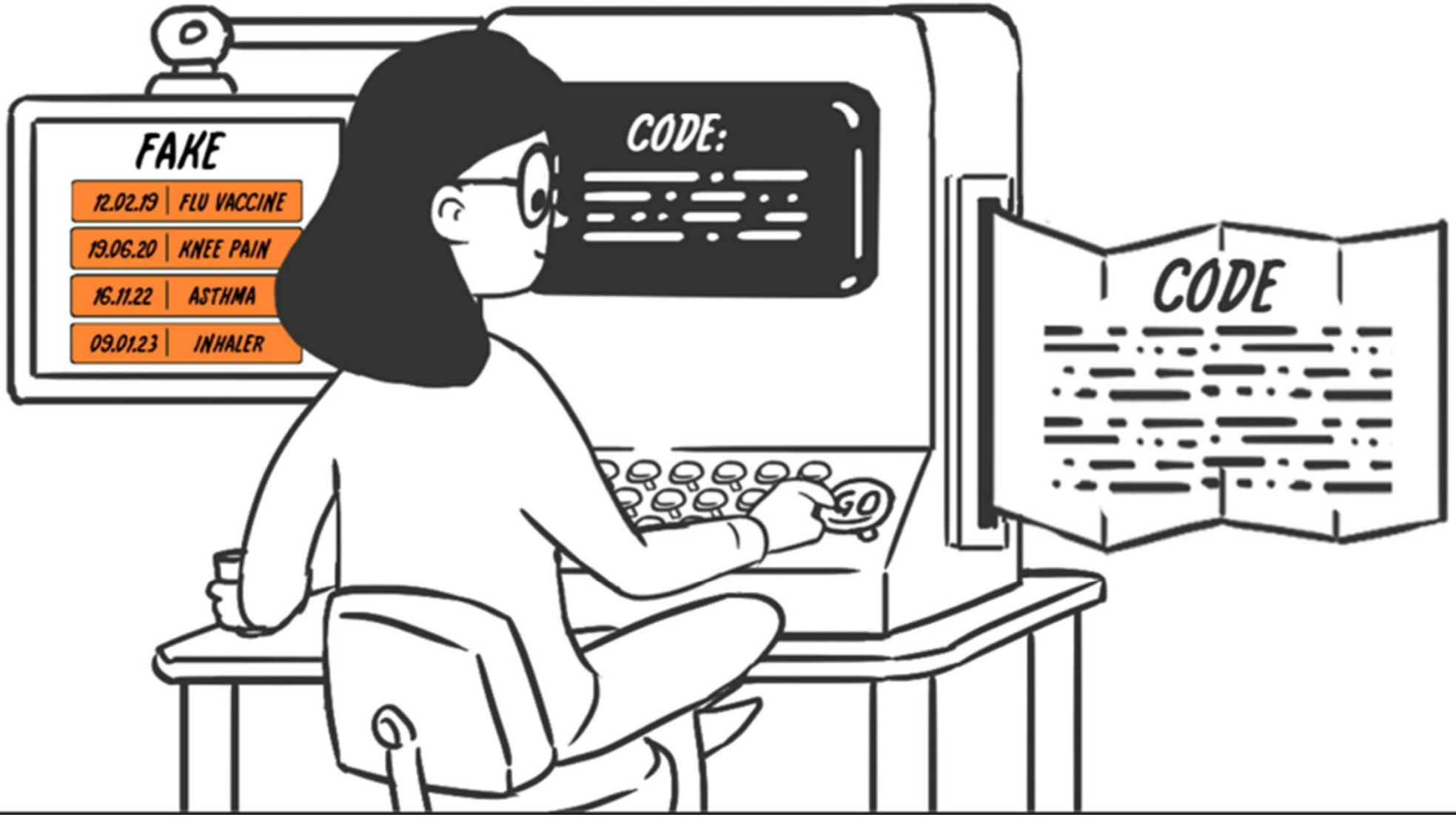

IT SOUNDS LIKE ASTHMA

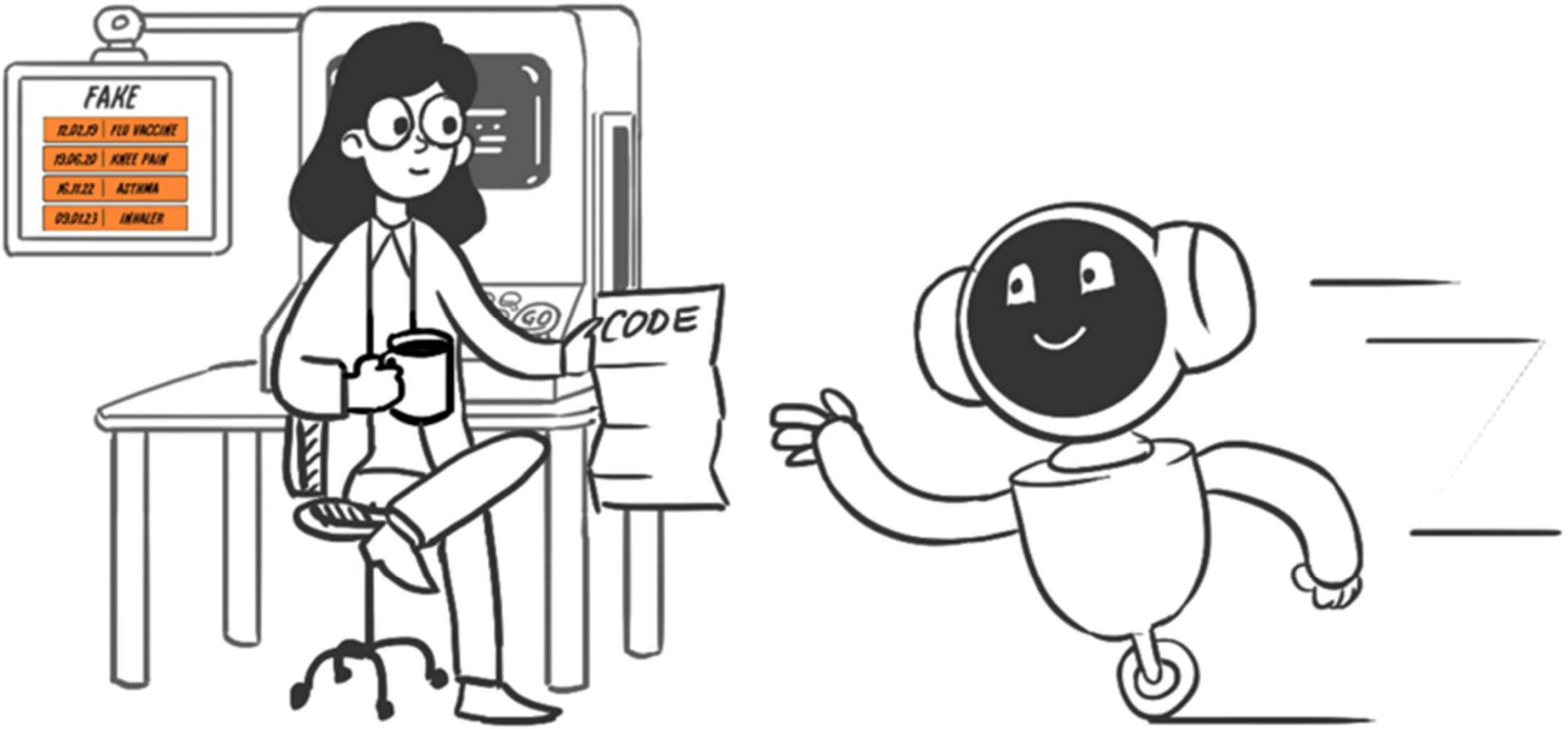

12.02.19 | FLU VACCINE
19.06.20 | KNEE PAIN
TODAY | ASTHMA

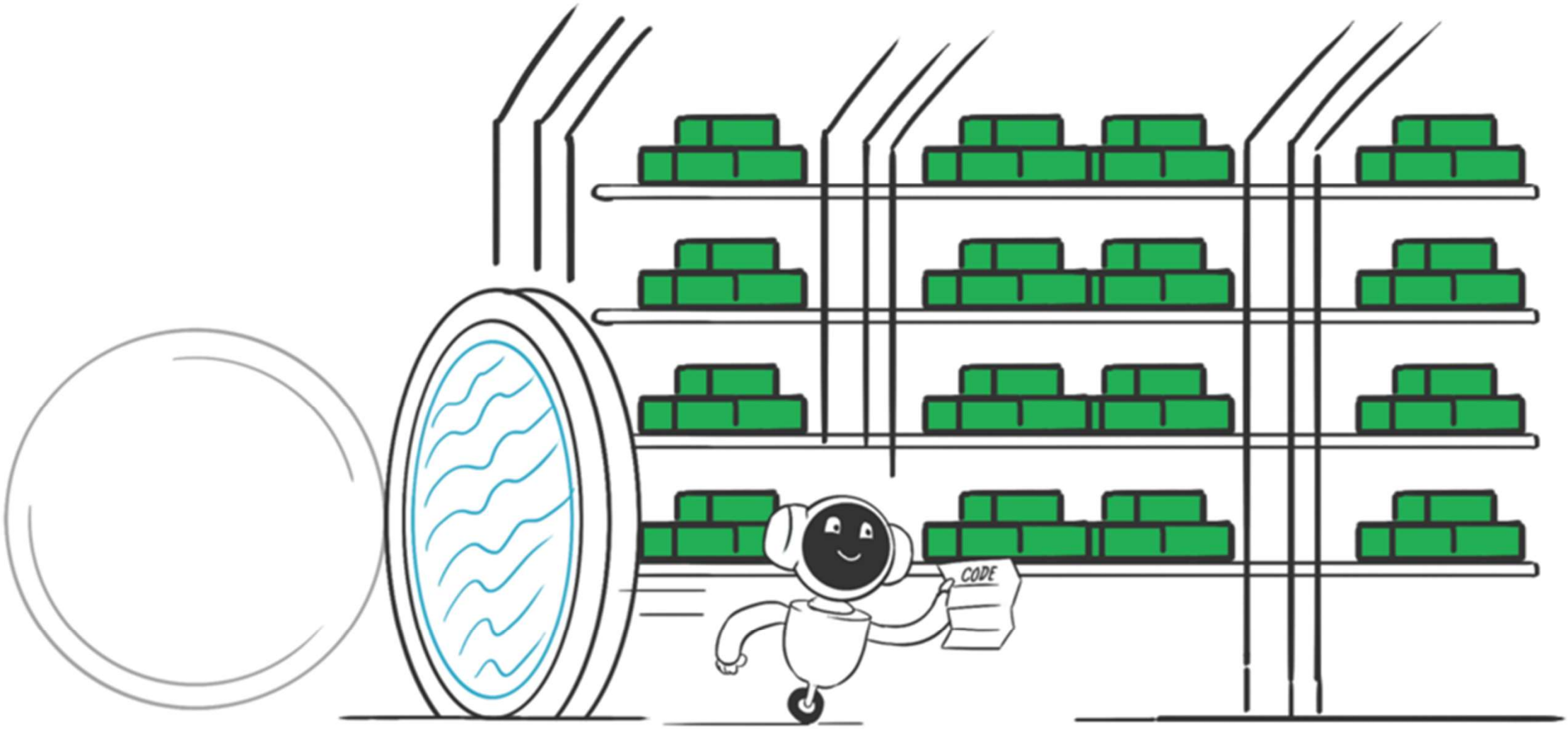

CODE:
19.06.20 | KNEE PAIN
16.11.22 | ASTHMA
TODAY | INHALER
GO

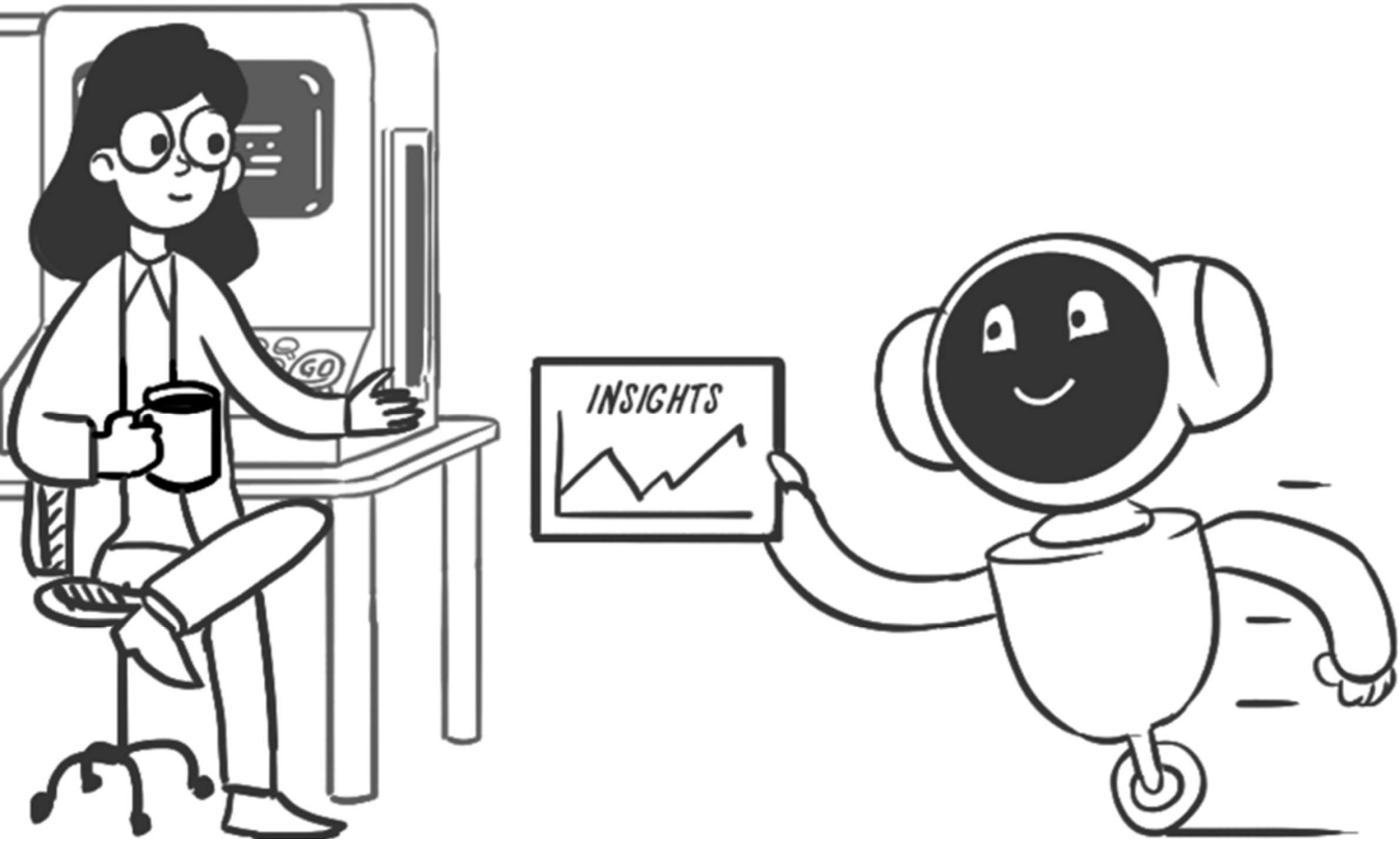

CODE:
GO


CODE:
GO
FAKE


FAKE
12.02.19 | FLU VACCINE
19.06.20 | KNEE PAIN
16.11.22 | ASTHMA
09.01.23 | INHALER
CODE:
GO
CODE

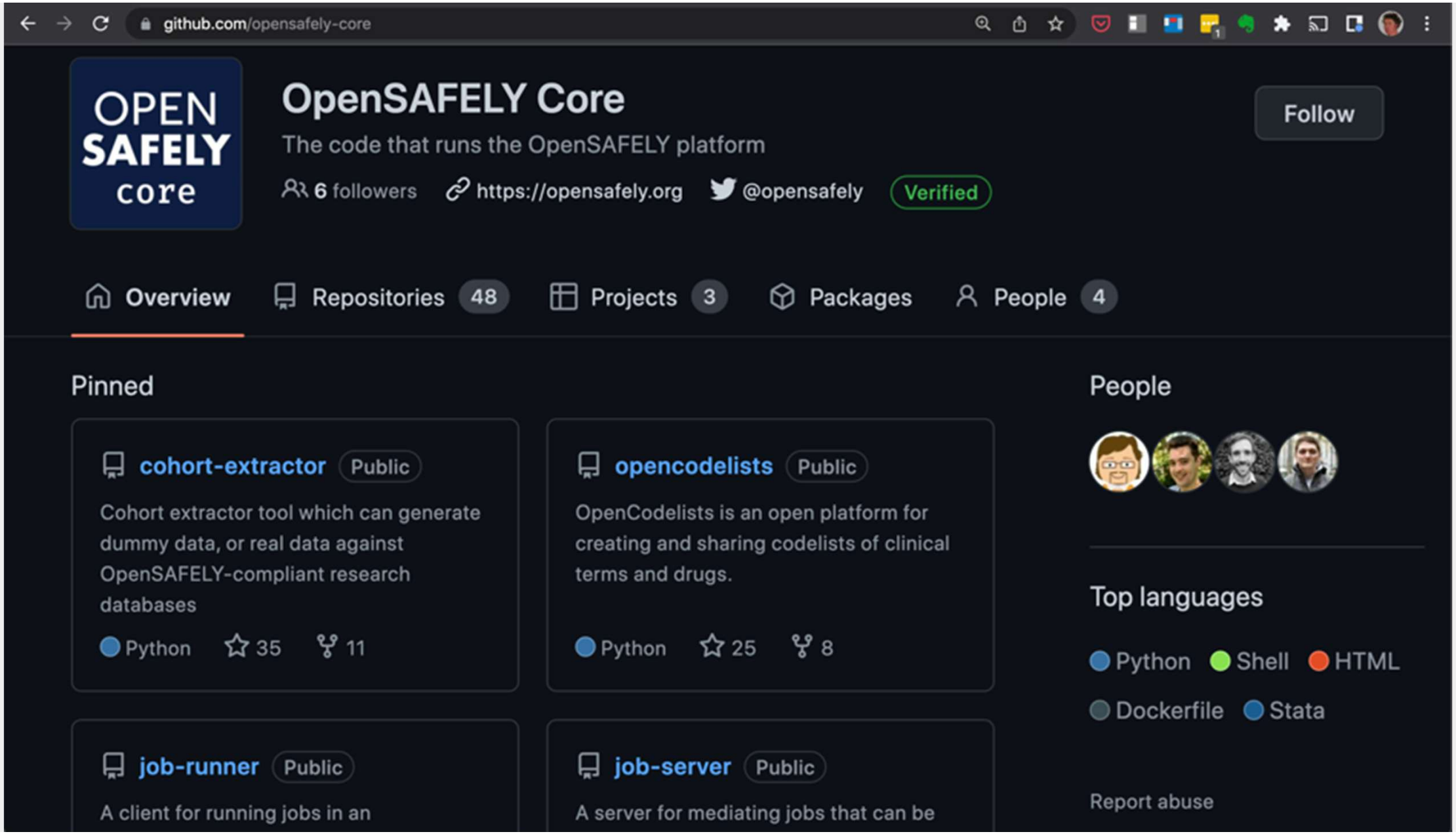

FAKE
12.01.23 | FLU VACCINE
13.06.20 | KNEE PAIN
16.11.22 | ASTHMA
09.01.23 | INHALER
CODE

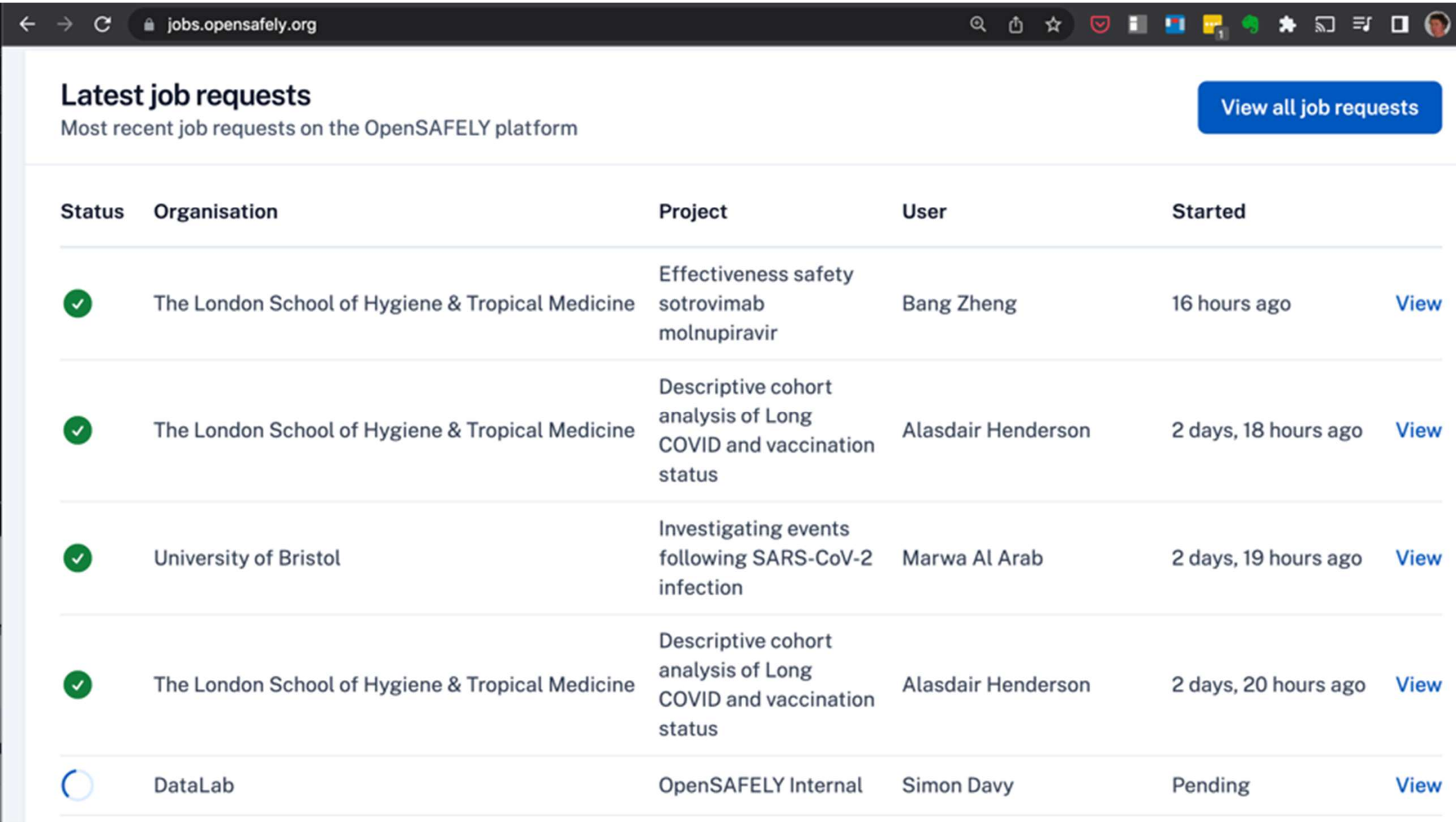

CODE


INSIGHTS

Trust

NHSE and GP as data controller
Records don’t move
Users work on dummy data
Total, “code-level” transparency


github.com/opensafely-core
OPEN SAFELY core
OpenSAFELY Core
The code that runs the OpenSAFELY platform
6 followers
https://opensafely.org
@opensafely
Verified
Follow
Overview
Repositories 48
Projects 3
Packages
People 4
Pinned
cohort-extractor Public
Cohort extractor tool which can generate dummy data, or real data against OpenSAFELY-compliant research databases
Python 35 11
opencodelists Public
OpenCodelists is an open platform for creating and sharing codelists of clinical terms and drugs.
Python 25 8
job-runner Public
A client for running jobs in an
job-server Public
A server for mediating jobs that can be
People
Top languages
Python Shell HTML
Dockerfile Stata
Report abuse


jobs.opensafely.org
Latest job requests
Most recent job requests on the OpenSAFELY platform
View all job requests
Status
Organisation
Project
User
Started
The London School of Hygiene & Tropical Medicine
Effectiveness safety sotrovimab molnupiravir
Bang Zheng
16 hours ago
View
The London School of Hygiene & Tropical Medicine
Descriptive cohort analysis of Long COVID and vaccination status
Alasdair Henderson
2 days, 18 hours ago
View
University of Bristol
Investigating events following SARS-CoV-2 infection
Marwa Al Arab
2 days, 19 hours ago
View
The London School of Hygiene & Tropical Medicine
Descriptive cohort analysis of Long COVID and vaccination status
Alasdair Henderson
2 days, 20 hours ago
View
DataLab
OpenSAFELY Internal
Simon Davy
Pending
View

jobs.opensafely.org/lshtm/effectiveness-safety-sotrovimab-molnupiravir/sotrovimab_molnupiravir/175...

OpenSAFELY Jobs

Event Log Status Login

The London School of Hygiene & Tropical Medicine > Effectiveness safety sotrovimab molnupiravir > sotrovimab_molnupiravir > 17508

Job request: 17508

sotrovimab_molnupiravir </> tfux2p2qqc5kstep

View repo View project.yaml

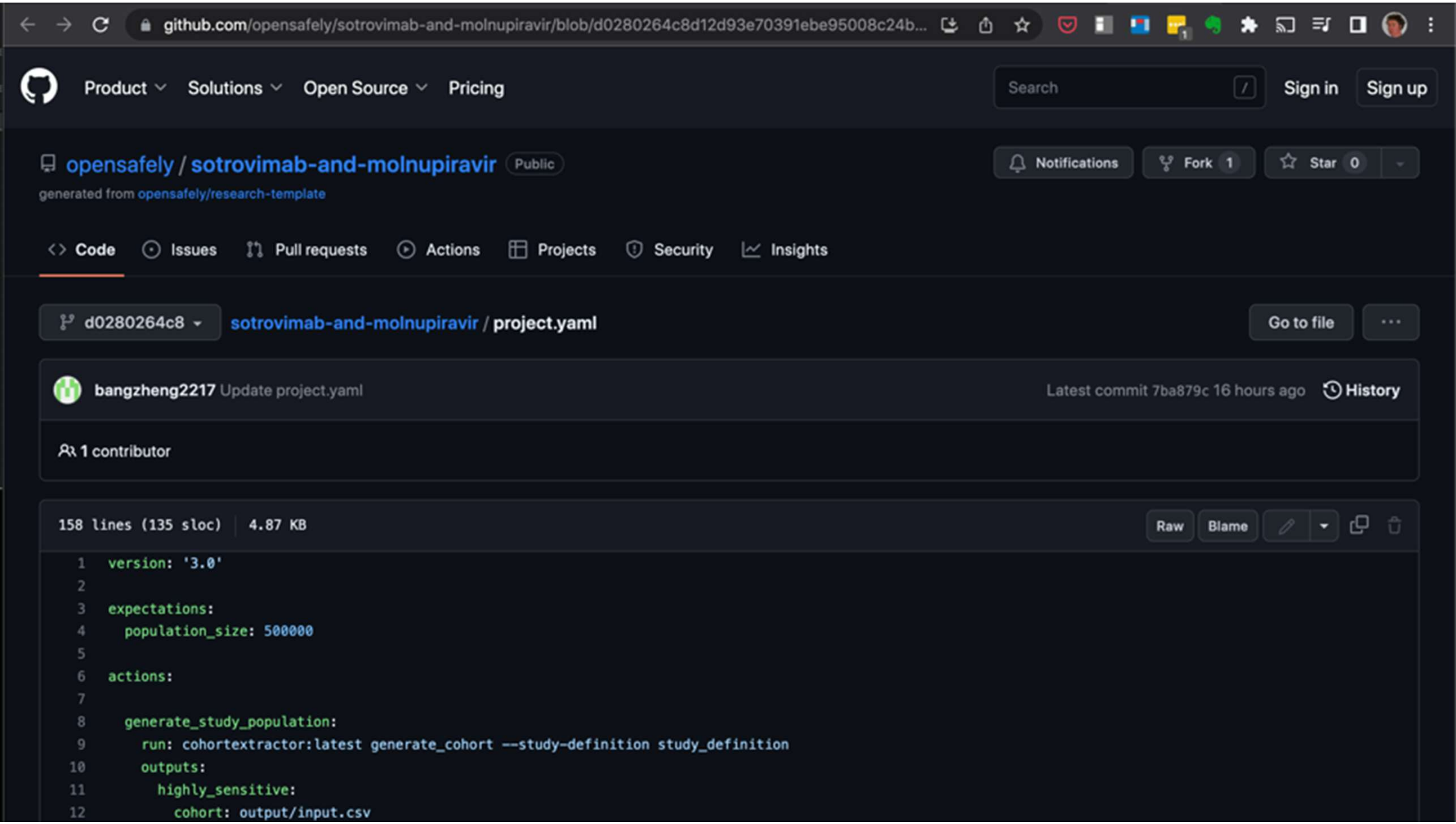
This page shows the technical details of what happened when authorised researcher **Bang Zheng** requested one or more actions to be run against real patient data in the project, within a secure environment.

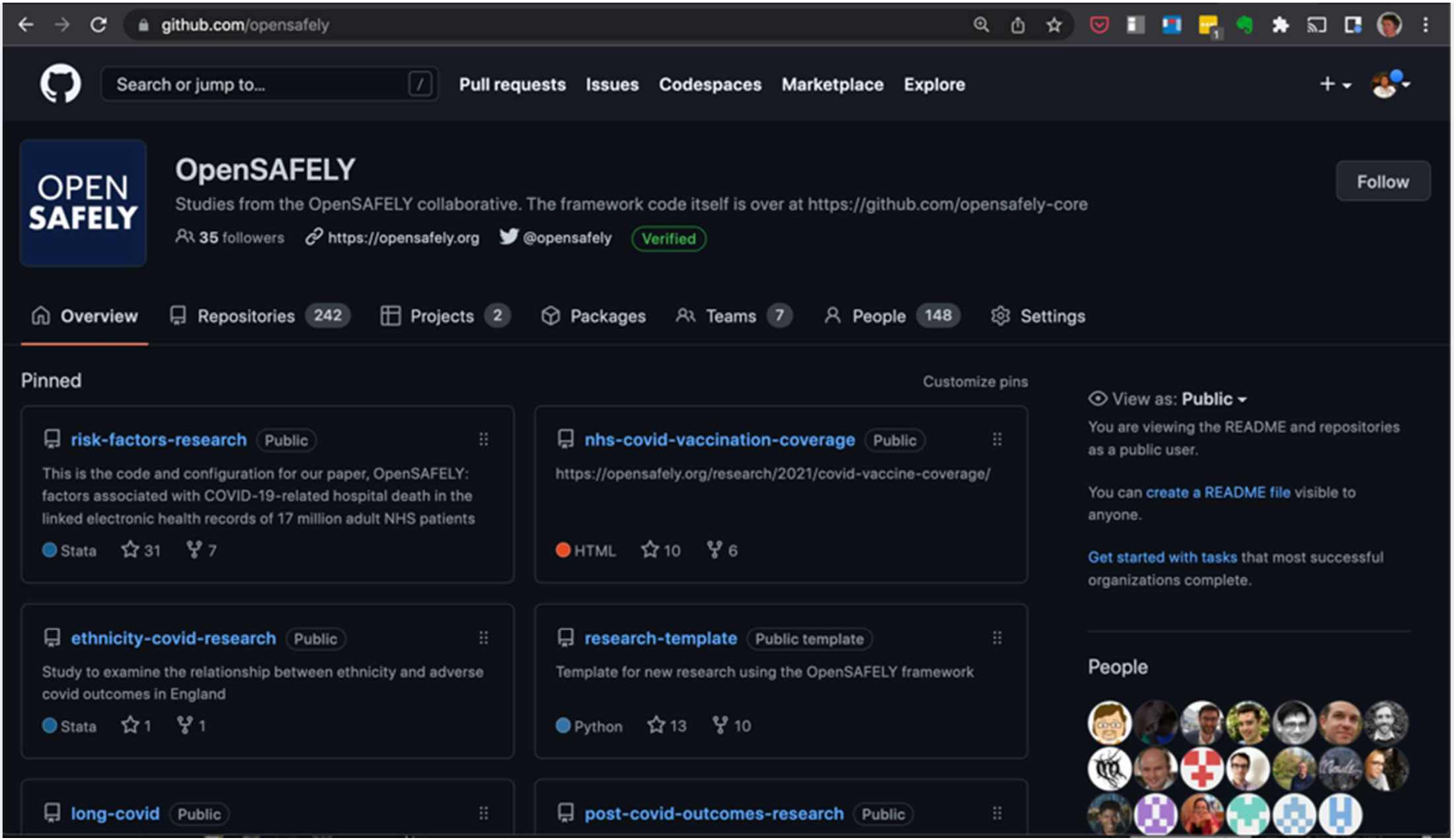
By cross-referencing the indicated Requested Actions with the Pipeline section below, you can infer what security level various outputs were written to. Outputs marked as `highly_sensitive` can never be viewed directly by a researcher; they can only request that code runs against them. Outputs

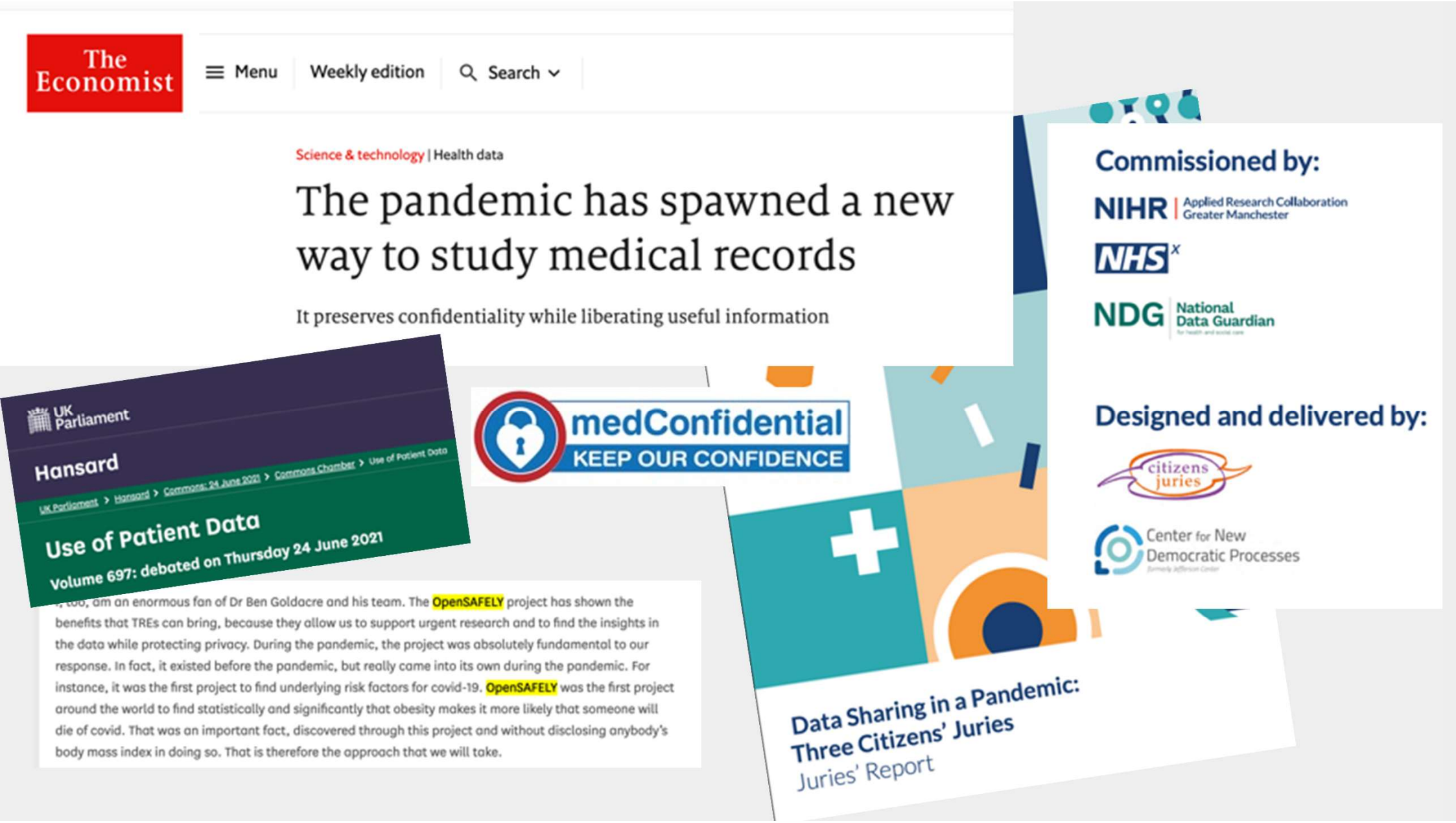
Timeline

Created:
16 hours ago

Started:
16 hours ago

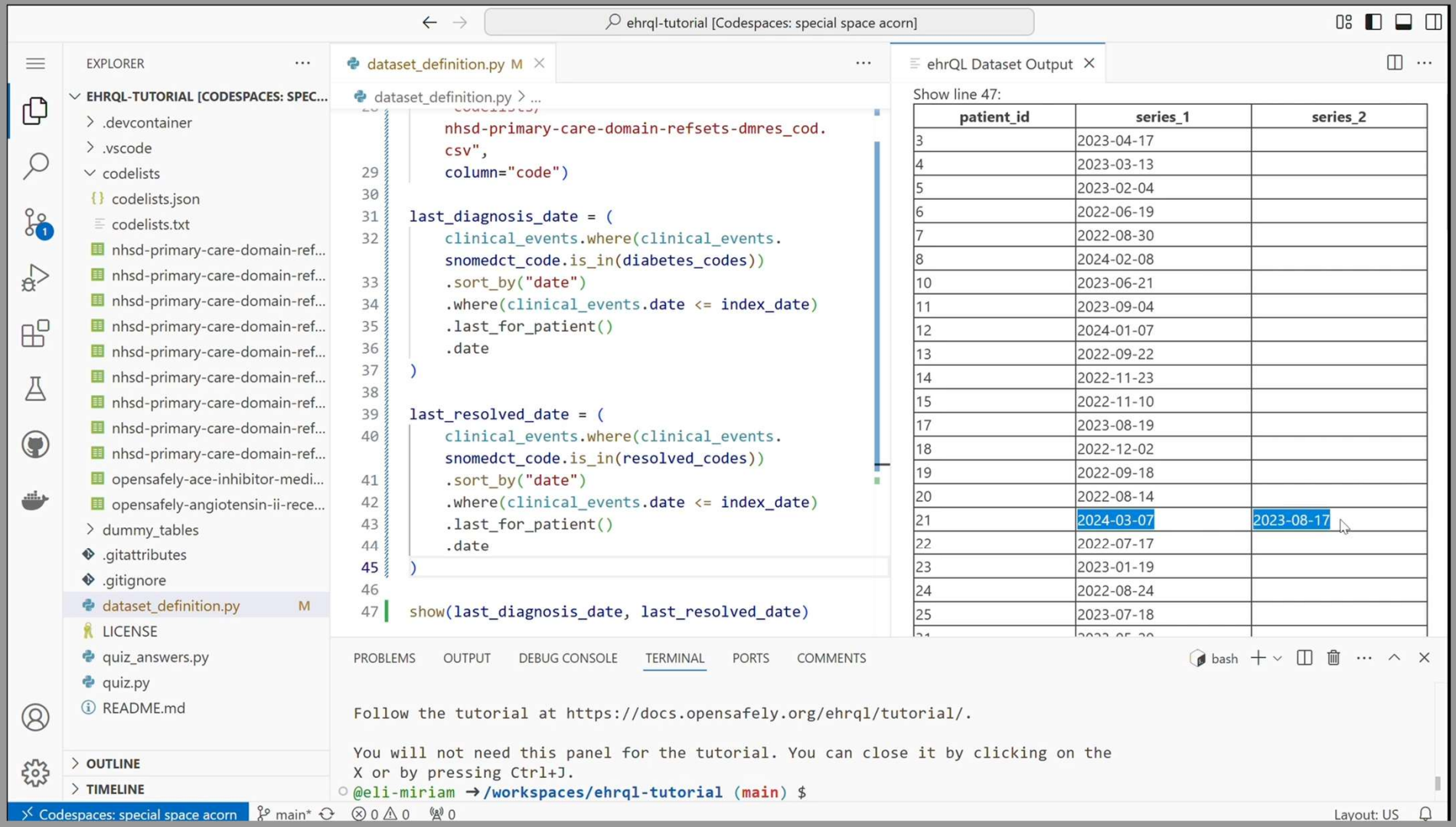

github.com/opensafely/sotrovimab-and-molnupiravir/blob/d0280264c8d12d93e70391ebe95008c24b...
Product
Solutions
Open Source
Pricing
Search
Sign in
Sign up
opensafely / sotrovimab-and-molnupiravir
Public
generated from opensafely/research-template
Notifications
Fork 1
Star 0
Code
Issues
Pull requests
Actions
Projects
Security
Insights
d0280264c8
sotrovimab-and-molnupiravir / project.yaml
Go to file
bangzheng2217 Update project.yaml
Latest commit 7ba879c 16 hours ago
History
1 contributor
158 lines (135 sloc)
4.87 KB
Raw
Blame
version: '3.0'
expectations:
population_size: 500000
actions:
generate_study_population:
run: cohortextractor:latest generate_cohort --study-definition study_definition
outputs:
highly_sensitive:
cohort: output/input.csv

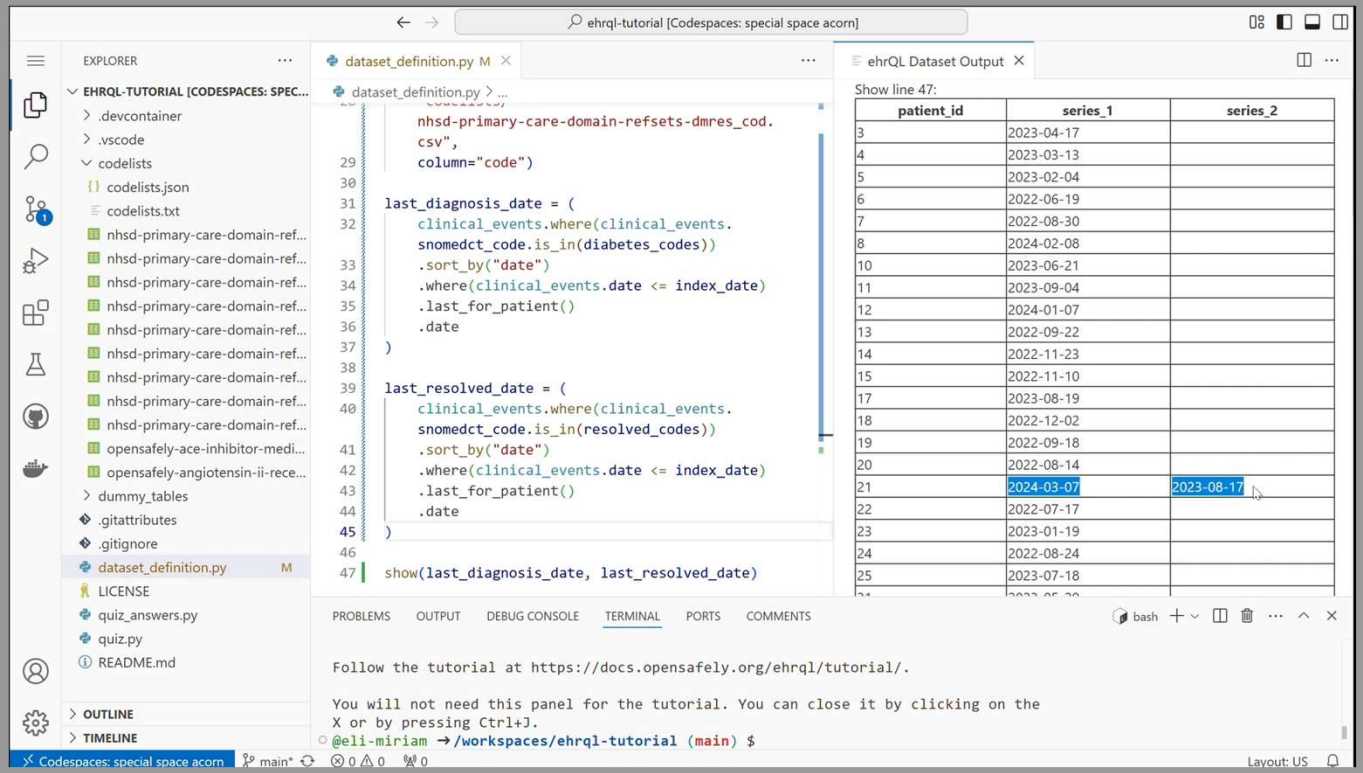

github.com/opensafely
Search or jump to...
Pull requests
Issues
Codespaces
Marketplace
Explore
OPEN SAFELY
OpenSAFELY
Studies from the OpenSAFELY collaborative. The framework code itself is over at https://github.com/opensafely-core
35 followers
https://opensafely.org
@opensafely
Verified
Follow
Overview
Repositories 242
Projects 2
Packages
Teams 7
People 148
Settings
Pinned
Customize pins
risk-factors-research Public
This is the code and configuration for our paper, OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients
Stata 31 7
nhs-covid-vaccination-coverage Public
https://opensafely.org/research/2021/covid-vaccine-coverage/
HTML 10 6
ethnicity-covid-research Public
Study to examine the relationship between ethnicity and adverse covid outcomes in England
Stata 1 1
research-template Public template
Template for new research using the OpenSAFELY framework
Python 13 10
long-covid Public
post-covid-outcomes-research Public
View as: Public
You are viewing the README and repositories as a public user.
You can create a README file visible to anyone.
Get started with tasks that most successful organizations complete.
People


The Economist
Menu
Weekly edition
Search
Science & technology | Health data
The pandemic has spawned a new way to study medical records
It preserves confidentiality while liberating useful information
UK Parliament
Hansard
UK Parliament > Hansard > Commons: 24 June 2021 > Commons Chamber > Use of Patient Data
Use of Patient Data
Volume 697: debated on Thursday 24 June 2021
, too, am an enormous fan of Dr Ben Goldacre and his team. The OpenSAFELY project has shown the benefits that TREs can bring, because they allow us to support urgent research and to find the insights in the data while protecting privacy. During the pandemic, the project was absolutely fundamental to our response. In fact, it existed before the pandemic, but really came into its own during the pandemic. For instance, it was the first project to find underlying risk factors for covid-19. OpenSAFELY was the first project around the world to find statistically and significantly that obesity makes it more likely that someone will die of covid. That was an important fact, discovered through this project and without disclosing anybody's body mass index in doing so. That is therefore the approach that we will take.
medConfidential
KEEP OUR CONFIDENCE
Data Sharing in a Pandemic:
Three Citizens' Juries
Juries' Report
Commissioned by:
NIHR | Applied Research Collaboration Greater Manchester
NHSX
NDG National Data Guardian
Designed and delivered by:
citizens juries
Center for New Democratic Processes


OpenSAFELY
Research
opensafely.org/research

What does it look like to a user?


ehrql-tutorial [Codespaces: special space acorn]
EXPLORER
EHRQL-TUTORIAL [CODESPACES: SPEC...
.devcontainer
.vscode
codelists
codelists.json
codelists.txt
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
nhsd-primary-care-domain-ref...
opensafely-ace-inhibitor-medi...
opensafely-angiotensin-ii-rece...
dummy_tables
.gitattributes
.gitignore
dataset_definition.py M
LICENSE
quiz_answers.py
quiz.py
README.md
OUTLINE
TIMELINE
dataset_definition.py M
dataset_definition.py > ...
nhsd-primary-care-domain-refsets-dmres_cod.
csv",
column="code")
last_diagnosis_date = (
clinical_events.where(clinical_events.
snomedct_code.is_in(diabetes_codes))
.sort_by("date")
.where(clinical_events.date <= index_date)
.last_for_patient()
.date
)
last_resolved_date = (
clinical_events.where(clinical_events.
snomedct_code.is_in(resolved_codes))
.sort_by("date")
.where(clinical_events.date <= index_date)
.last_for_patient()
.date
)
show(last_diagnosis_date, last_resolved_date)
ehrQL Dataset Output
Show line 47:
patient_id series_1 series_2
3 2023-04-17
4 2023-03-13
5 2023-02-04
6 2022-06-19
7 2022-08-30
8 2024-02-08
10 2023-06-21
11 2023-09-04
12 2024-01-07
13 2022-09-22
14 2022-11-23
15 2022-11-10
17 2023-08-19
18 2022-12-02
19 2022-09-18
20 2022-08-14
21 2024-03-07 2023-08-17
22 2022-07-17
23 2023-01-19
24 2022-08-24
25 2023-07-18
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS COMMENTS
bash
Follow the tutorial at https://docs.opensafely.org/ehrql/tutorial/.
You will not need this panel for the tutorial. You can close it by clicking on the
X or by pressing Ctrl+J.
@eli-miriam → /workspaces/ehrql-tutorial (main) $
Codespaces: special space acorn
main*
Layout: US



R, Stata, Python etc

Pick a backend to run your Jobs in:

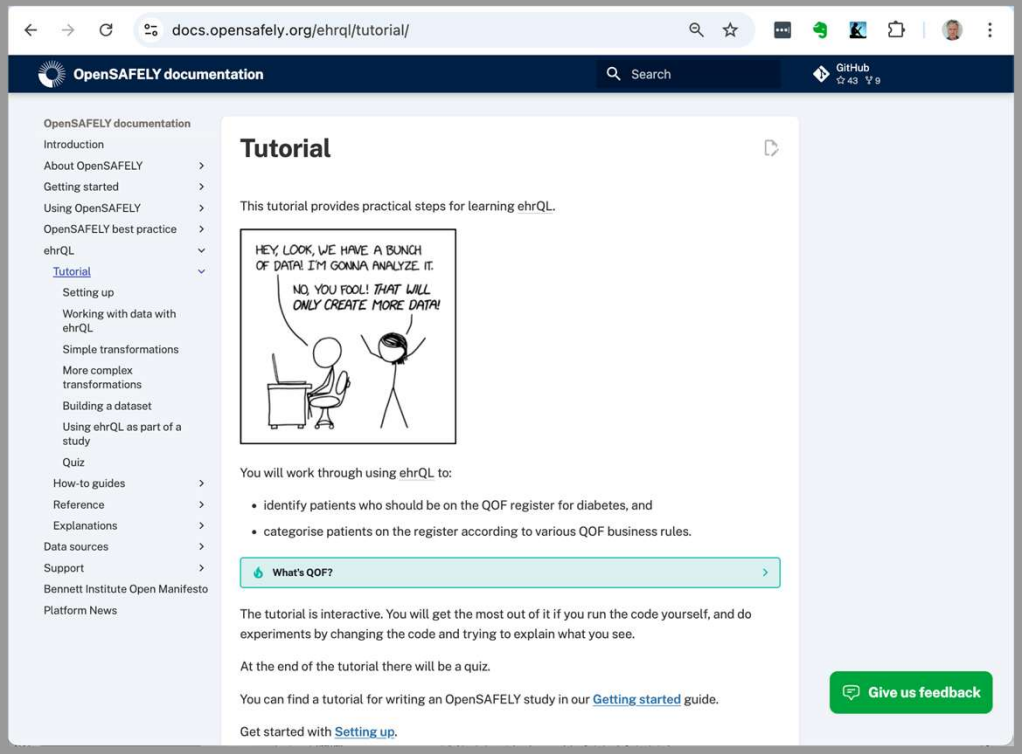
- ○ EMIS
- ◉ TPP

Pick one or more actions to run:

Select	Action	Previously run	Status
Run	`generate_dataset`	✓	
Run	`analyse_dataset`	✓	Needs
Run	`run_all`		Needs

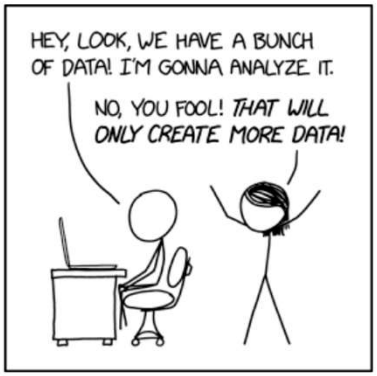
Submit

docs.opensafely.org



Tutorial

This tutorial provides practical steps for learning ehrQL.



You will work through using ehrQL to:

- identify patients who should be on the QOF register for diabetes, and
- categorise patients on the register according to various QOF business rules.

What's QOF?

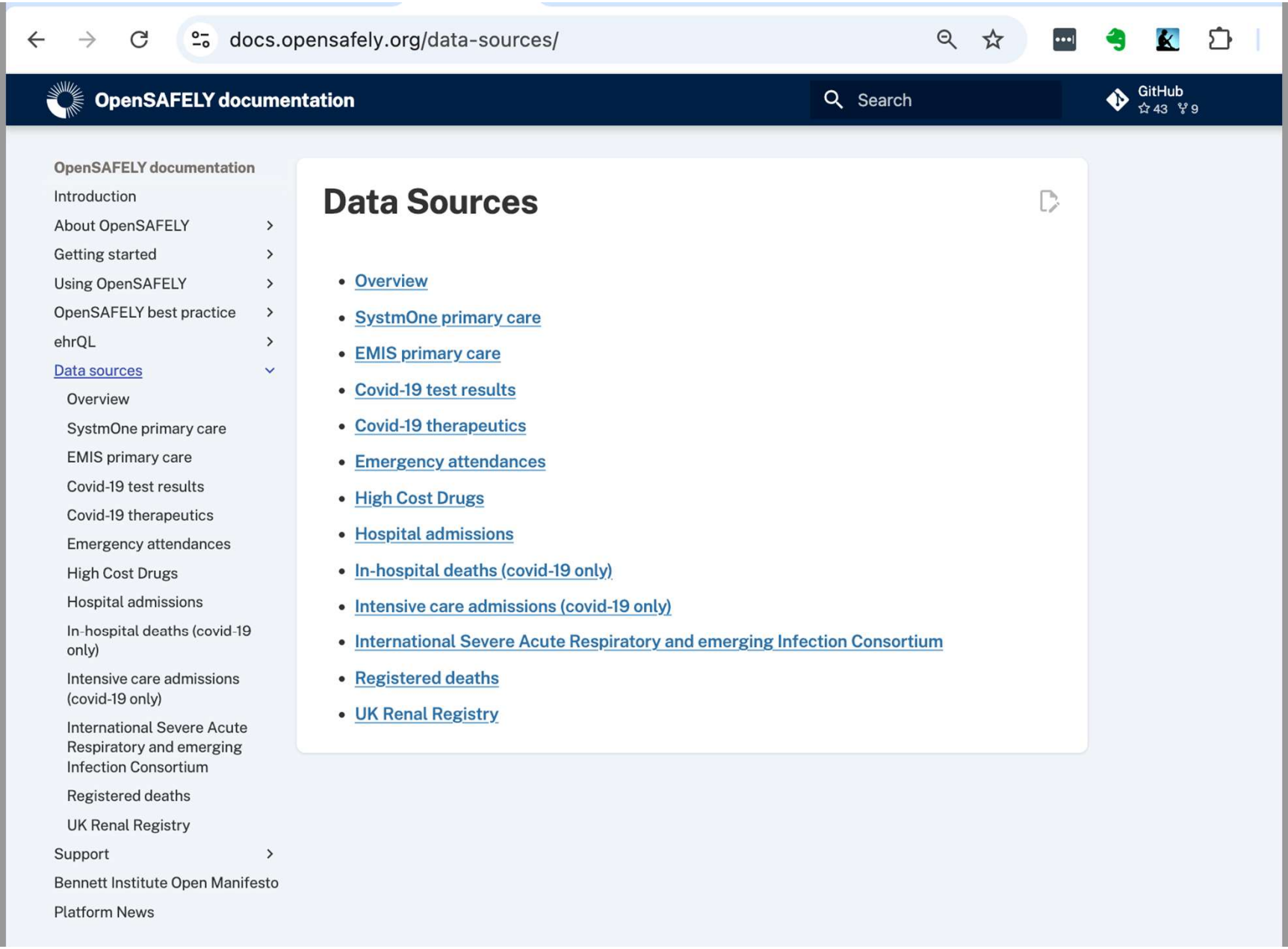
The tutorial is interactive. You will get the most out of it if you run the code yourself, and do experiments by changing the code and trying to explain what you see.

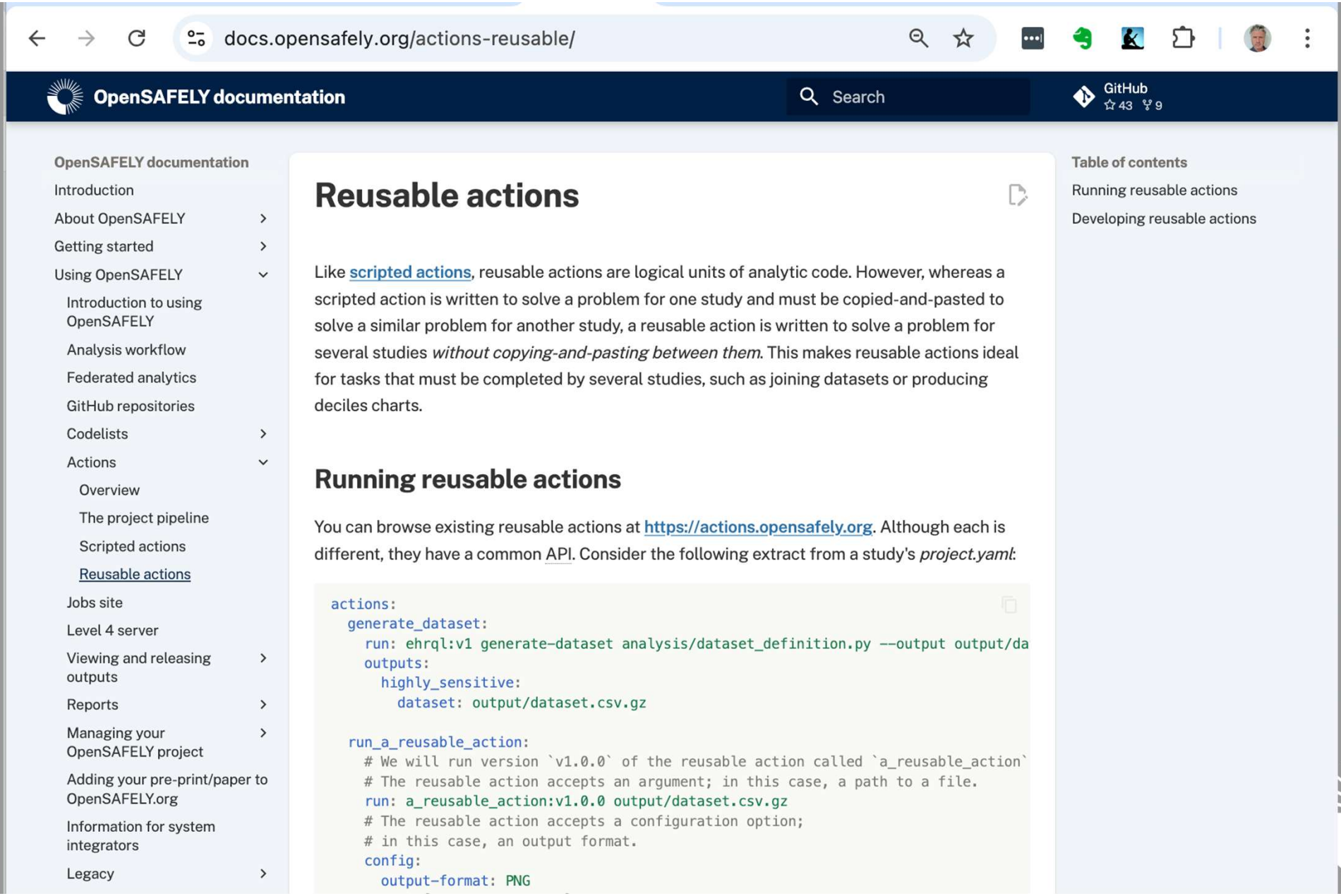
At the end of the tutorial there will be a quiz.

You can find a tutorial for writing an OpenSAFELY study in our Getting started guide.

Get started with Setting up.

Give us feedback


docs.opensafely.org/data-sources/
OpenSAFELY documentation
Search
GitHub
43
9
OpenSAFELY documentation
Introduction
About OpenSAFELY
Getting started
Using OpenSAFELY
OpenSAFELY best practice
ehrQL
Data sources
Overview
SystmOne primary care
EMIS primary care
Covid-19 test results
Covid-19 therapeutics
Emergency attendances
High Cost Drugs
Hospital admissions
In-hospital deaths (covid-19 only)
Intensive care admissions (covid-19 only)
International Severe Acute Respiratory and emerging Infection Consortium
Registered deaths
UK Renal Registry
Support
Bennett Institute Open Manifesto
Platform News
Data Sources
Overview
SystmOne primary care
EMIS primary care
Covid-19 test results
Covid-19 therapeutics
Emergency attendances
High Cost Drugs
Hospital admissions
In-hospital deaths (covid-19 only)
Intensive care admissions (covid-19 only)
International Severe Acute Respiratory and emerging Infection Consortium
Registered deaths
UK Renal Registry


docs.opensafely.org/actions-reusable/
OpenSAFELY documentation
Search
GitHub
43
9
OpenSAFELY documentation
Introduction
About OpenSAFELY
Getting started
Using OpenSAFELY
Introduction to using OpenSAFELY
Analysis workflow
Federated analytics
GitHub repositories
Codelists
Actions
Overview
The project pipeline
Scripted actions
Reusable actions
Jobs site
Level 4 server
Viewing and releasing outputs
Reports
Managing your OpenSAFELY project
Adding your pre-print/paper to OpenSAFELY.org
Information for system integrators
Legacy
Reusable actions
Like scripted actions, reusable actions are logical units of analytic code. However, whereas a scripted action is written to solve a problem for one study and must be copied-and-pasted to solve a similar problem for another study, a reusable action is written to solve a problem for several studies without copying-and-pasting between them. This makes reusable actions ideal for tasks that must be completed by several studies, such as joining datasets or producing deciles charts.
Running reusable actions
You can browse existing reusable actions at https://actions.opensafely.org. Although each is different, they have a common API. Consider the following extract from a study's project.yaml:
actions:
generate_dataset:
run: ehrql:v1 generate-dataset analysis/dataset_definition.py --output output/da
outputs:
highly_sensitive:
dataset: output/dataset.csv.gz
run_a_reusable_action:
We will run version `v1.0.0` of the reusable action called `a_reusable_action`
The reusable action accepts an argument; in this case, a path to a file.
run: a_reusable_action:v1.0.0 output/dataset.csv.gz
The reusable action accepts a configuration option;
in this case, an output format.
config:
output-format: PNG
Table of contents
Running reusable actions
Developing reusable actions

OpenSAFELY co-pilot programme: assisting users on their OpenSAFELY journey

Posted: 5 Aug 2021
Written by: Amelia Green
Categories: OpenSAFELY

Self-service data preparation tools

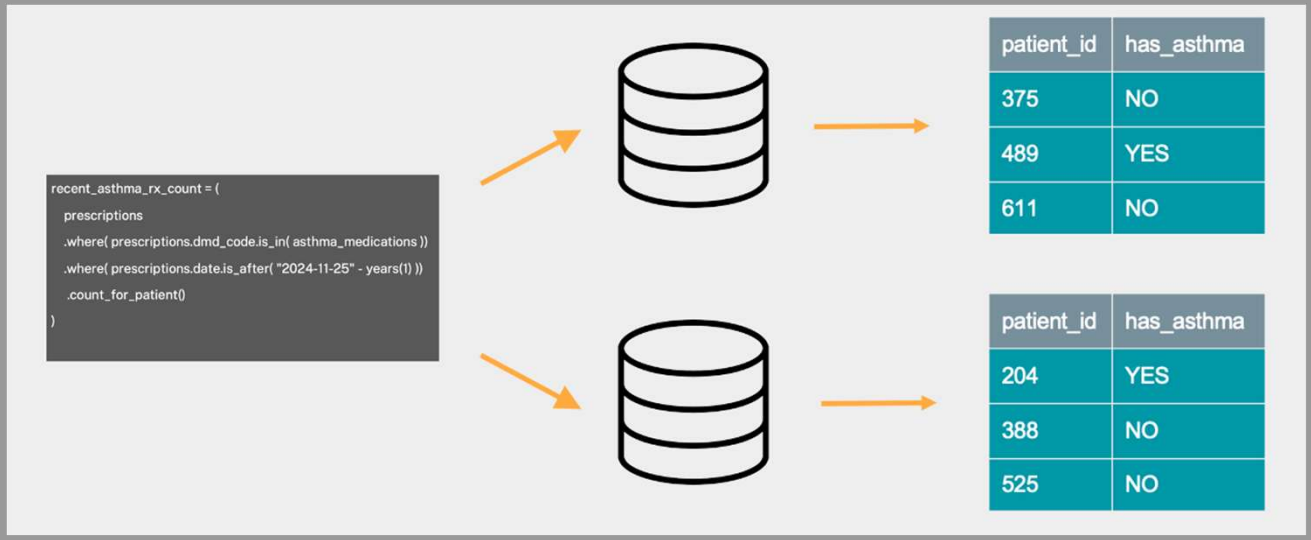
ehrQL

```
recent_asthma_rx_count = (
  prescriptions
  .where( prescriptions.dmd_code.is_in( asthma_medications ))
  .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
   .count_for_patient()
)
```

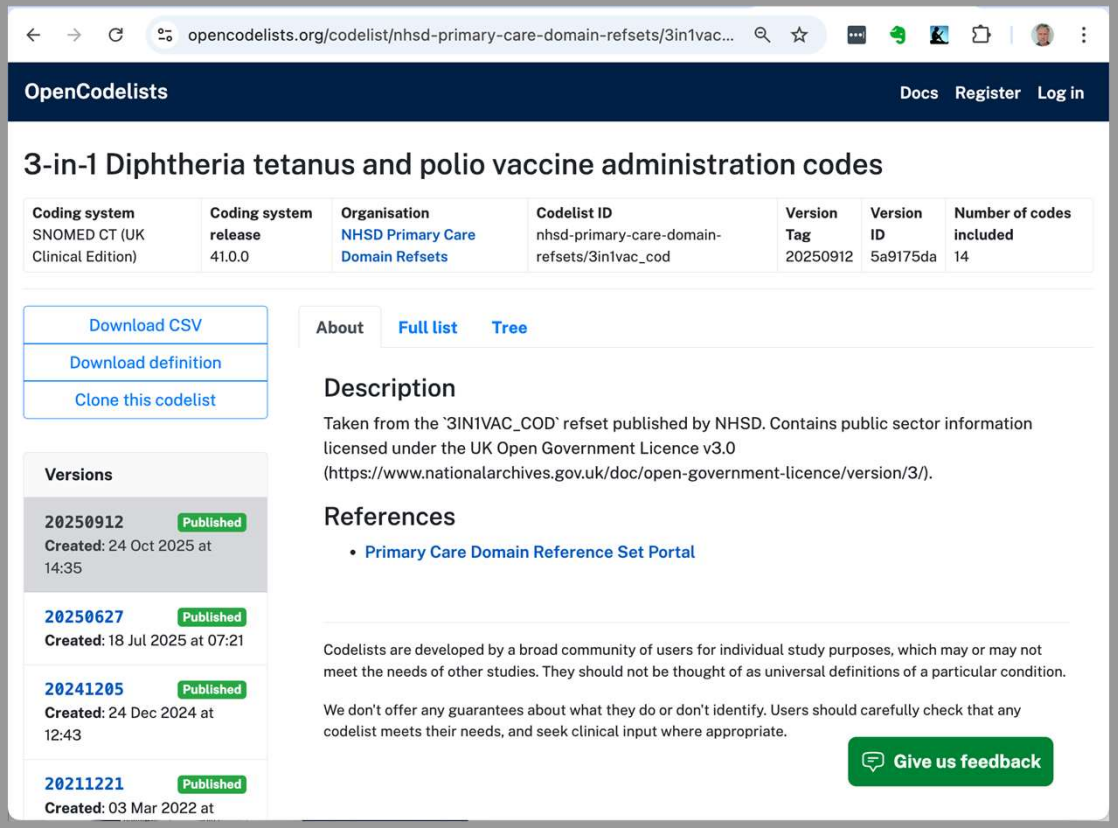
Every user can see every other user's work: understand it, re-use it, adapt it, and improve it.

Efficiency

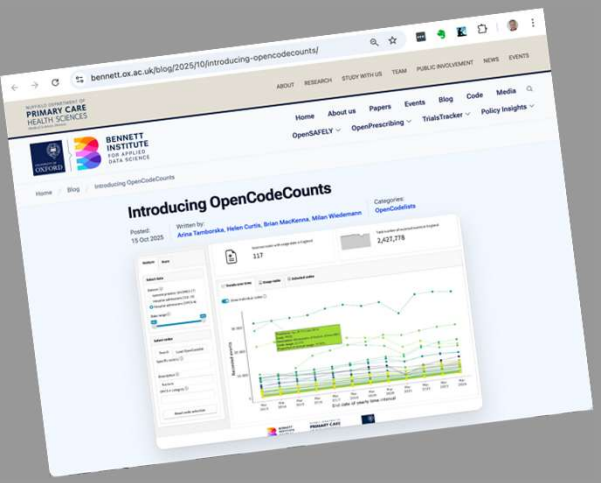
Standardised code (“ehrQL”)



OpenCodeLists.org



OpenCodeCounts



Analyse | More

Select data

Dataset
- General practice (SNOMED CT)
- Hospital admissions (ICD-10)
- Hospital admissions (OPCS-4)

Date range: 2013 – 2024

Select codes

Search | Load OpenCodelist

Specific code(s)

Description: fracture

OPCS-4 category

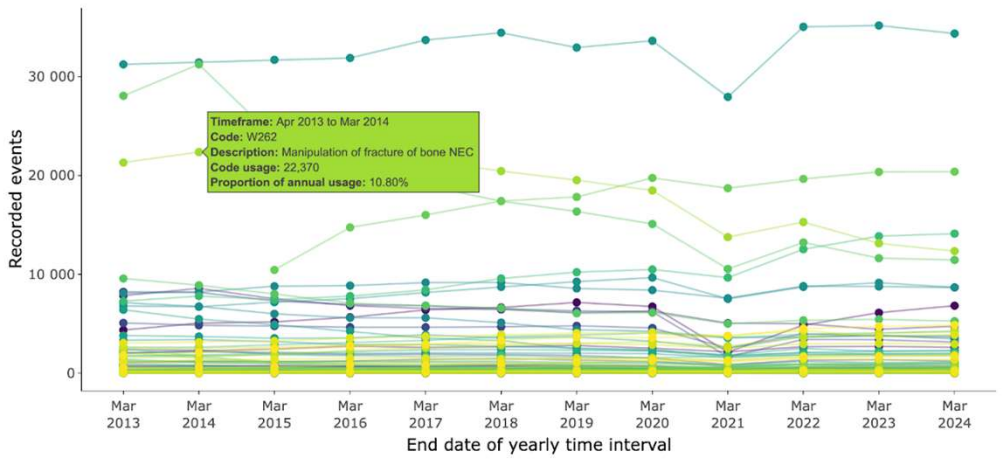
Reset code selection

Selected codes with usage data in England: 117

Total number of recorded events in England: 2,427,778

Trends over time | Usage table | Selected codes

Show individual codes



BENNETT INSTITUTE FOR APPLIED DATA SCIENCE
NUFFIELD DEPARTMENT OF PRIMARY CARE HEALTH SCIENCES
UNIVERSITY OF OXFORD

The data is locked down
Everything else is open

The full technical user manual	docs.opensafely.org
All the code for the platform	github.com/opensafely-core
All the code for every analysis	github.com/opensafely
Real-time logs of all code run	jobs.opensafely.org
All the approved projects	opensafely.org/approved-projects
All the outputs	opensafely.org/research

What's next?

Cautious expansion to non-COVID work
New features for users
Non-health data (OpenSAFELY-schools)
Non-UK data

What's next?

What's next?

non-COVID with NHSE

What's next?

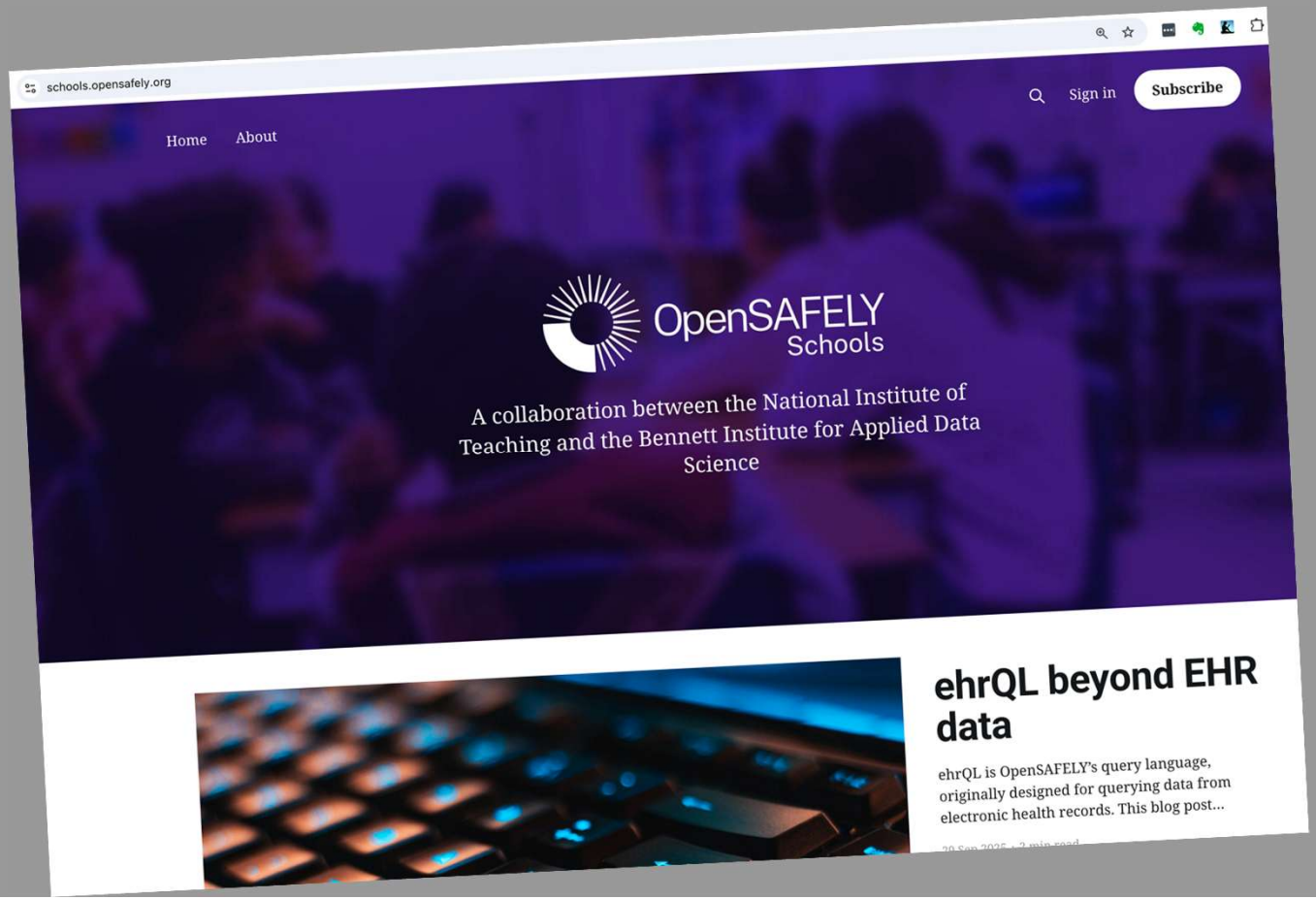
non-COVID with NHSE

What's next?

re-opening the EMIS service

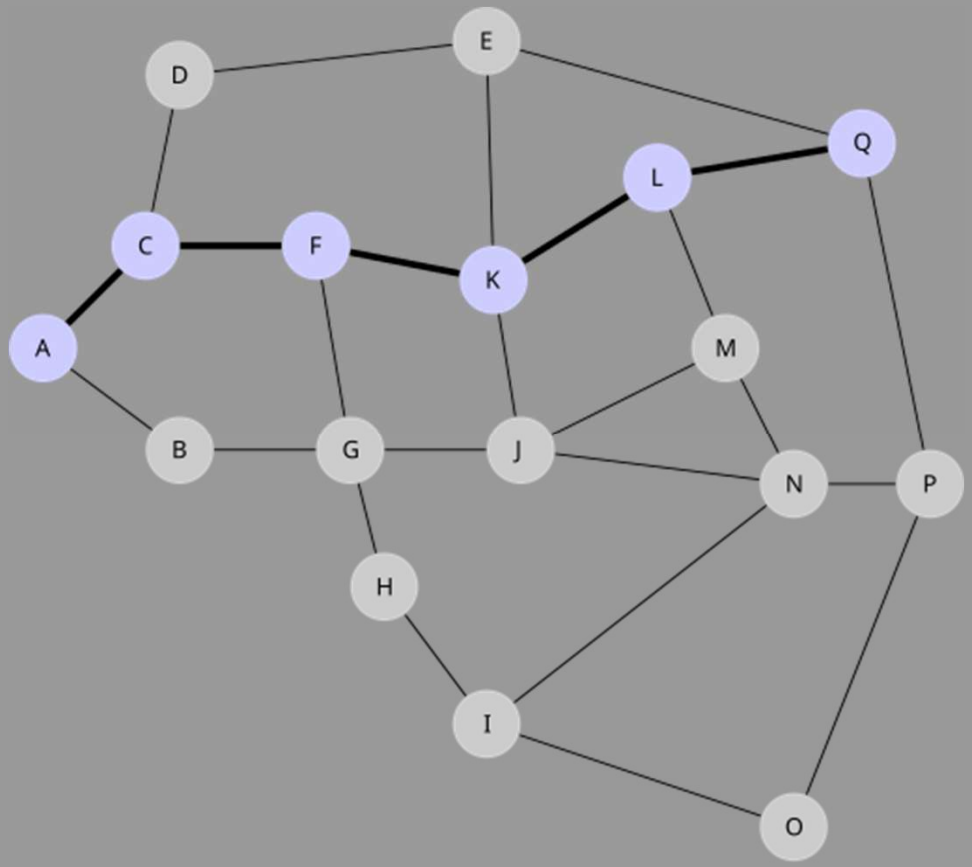
What's next?

OpenSAFELY-Schools



NATIONAL INSTITUTE of Teaching

What’s next?



Get involved

team@opensafely.org

Users

Use OpenSAFELY to access NHS data for research via NHS England.

Collaborate on the platform

Develop your own code modules in OpenSAFELY (EHR curation, automated output checking, automated workflows, more).

Use our tools in your data centre

Full OpenSAFELY installs; or just single modules (eg GP and EHR data preparation tools).

Get our help designing data ecosystems

We design efficient, accountable, end-to-end ecosystems, re-using the services, tools and teams that work.

NHS England OpenSAFELY GP Data Service

58 million	patients' records
181	projects
31	organisations
>100	completed outputs
42 days	from first shovel to first output
£10 million	spend to date, new funding to scale

Thank you!

Wellcome, RCGP, BMA, JGPITC, National Data Guardian, TPP, EMIS, NHS England, NIHR, MRC, LSHTM, Bristol, Manchester, and all our productive users!

OpenSAFELY

Working with EHR data is hard

“Raw” GP data is different to “analysis-ready” GP data

Raw GP data

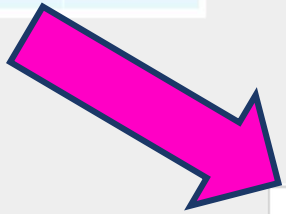
Patient ID	Event code	Associated variable	Event definition	Date, Time	Location
979384758	38341003		“Diagnosis of hypertension”	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	271649006	155	Blood Pressure systolic reading	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	VMP 318855006	28 tablets	Prescription for Enalapril	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA

Analysis-ready GP data

Patient ID	HT active in 2019	Cancer in prev 5 yrs	Total Oral Morphine Equivalent Prescribed in 2022
923459023490	Yes	No	2,200mg
239047293489089	No	No	0mg
23098409283490	No	Yes	120mg

Converting raw GP data into analysis-ready GP data is hard

Patient ID	Event code	Associated variable	Event definition	Date, Time	Location
979384758	38341003		"Diagnosis of hypertension"	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	271649006	155	Blood Pressure systolic reading	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	VMP 318855006	28 tablets	Prescription for Enalapril	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA



?

Patient ID	HT active in 2019	Cancer in prev 5 yrs	Total Oral Morphine Equivalent Prescribed in 2022
923459023490	Yes	No	2,200mg
239047293489089	No	No	0mg
23098409283490	No	Yes	120mg

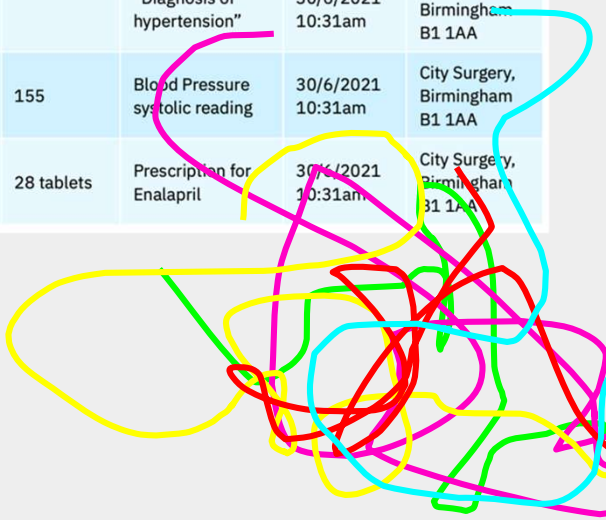
Converting raw GP data into analysis-ready GP data is hard and everyone does it differently!

Patient ID	Event code	Associated variable	Event definition	Date, Time	Location
979384758	38341003		"Diagnosis of hypertension"	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	271649006	155	Blood Pressure systolic reading	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	VMP 318855006	28 tablets	Prescription for Enalapril	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA

Patient ID	HT active in 2019	Cancer in prev 5 yrs	Total Oral Morphine Equivalent Prescribed in 2022
923450023490	Yes	No	2,200mg
23904729389089	No	No	0mg
230984092830490	No	Yes	120mg

Converting raw GP data into analysis-ready GP data is hard and everyone does it differently!

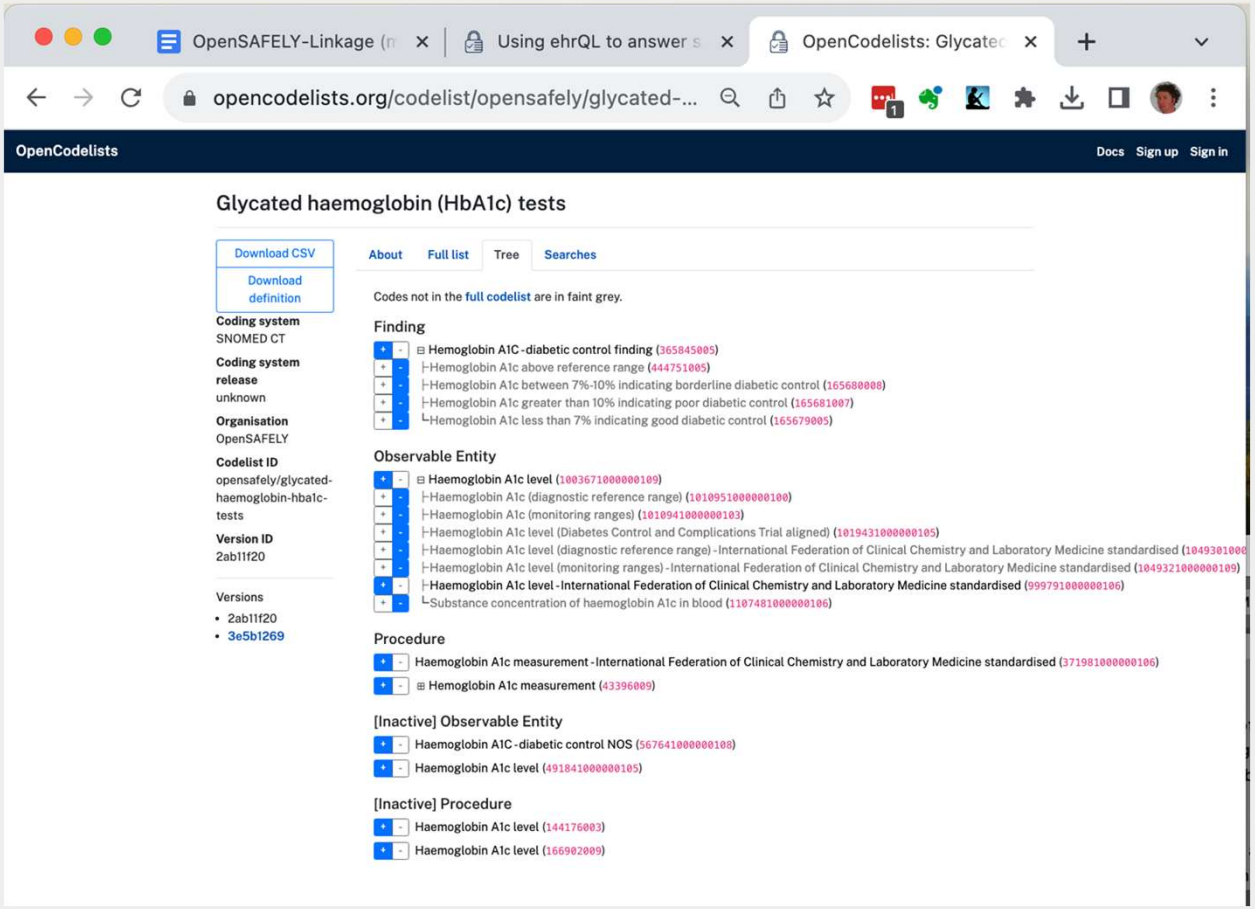
Patient ID	Event code	Associated variable	Event definition	Date, Time	Location
979384758	38341003		"Diagnosis of hypertension"	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	271649006	155	Blood Pressure systolic reading	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA
979384758	VMP 318855006	28 tablets	Prescription for Enalapril	30/6/2021 10:31am	City Surgery, Birmingham B1 1AA



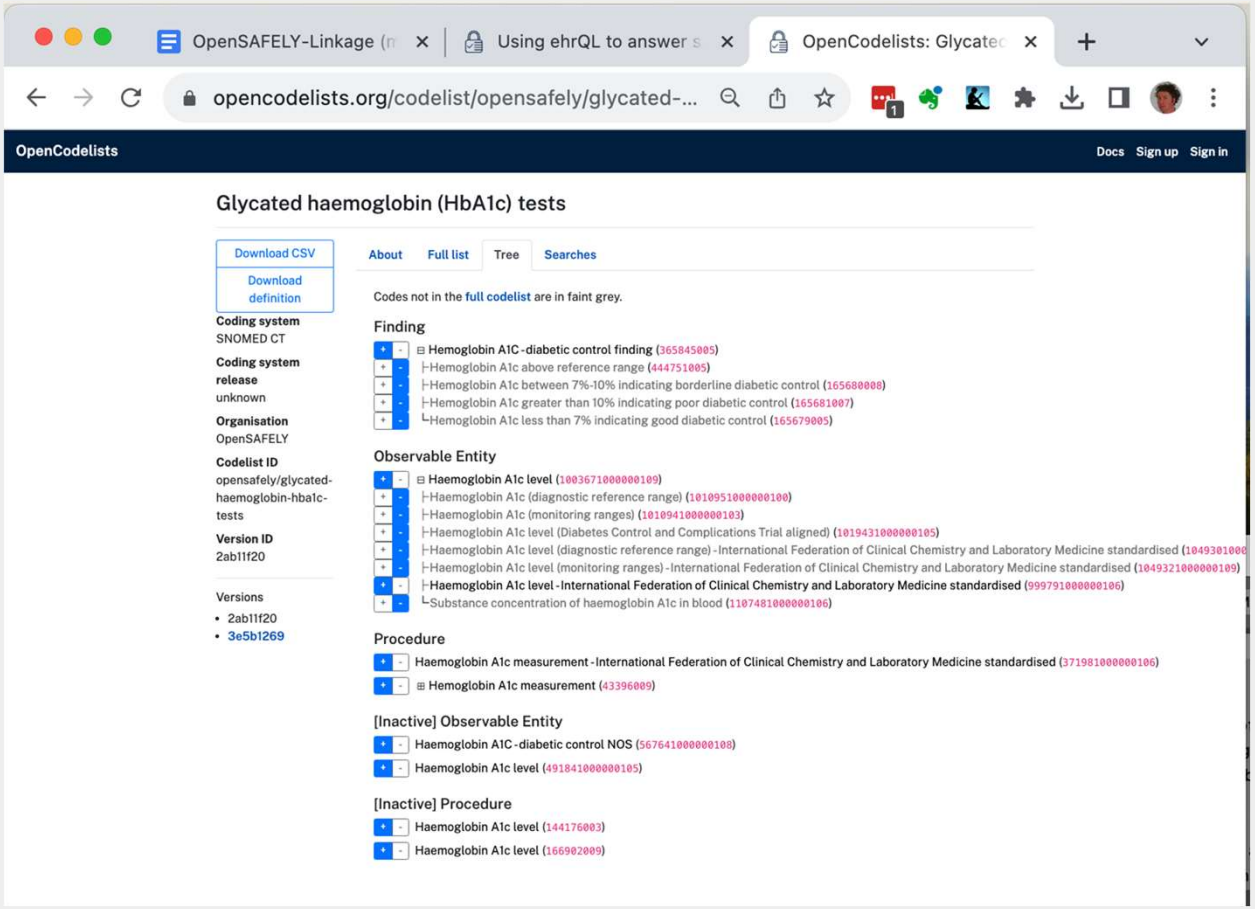
We wanted (and needed) to standardise EHR data management

Patient ID	HT active in 2019	Cancer in prev 5 yrs	Total Oral Morphine Equivalent Prescribed in 2022
923450023490	Yes	No	2,200mg
23904729389089	No	No	0mg
230984092830490	No	Yes	120mg

In OpenSAFELY, we made point and click tools to make, share, and assure “codelists”



In OpenSAFELY, we made point and click tools to make, share, and assure “codelists”



These don’t rely on OpenSAFELY.

You can easily use OpenCodelists yourselves, to make codelists for HES, GP data… that you already hold on your own machine…

In OpenSAFELY, we use a dynamic set of standard code and tools for EHR data preparation

Getting properties of an event matching some criteria

What is the code of the first/last clinical event matching some criteria?

```
from ehrql import create_dataset, codelist_from_csv
from ehrql.tables.beta.core import clinical_events, patients

asthma_codelist = codelist_from_csv("XXX", column="YYY")

dataset = create_dataset()
dataset.first_asthma_diagnosis_date = clinical_events.where(
        clinical_events.snomedct_code.is_in(asthma_codelist)
).where(
        clinical_events.date.is_on_or_after("2022-07-01")
).sort_by(
        clinical_events.date
).first_for_patient().snomedct_code
dataset.define_population(patients.exists_for_patient())
```

Performing arithmetic on numeric values of clinical events

Finding the mean observed value of clinical events matching some criteria

```
from ehrql import create_dataset, codelist_from_csv
from ehrql.tables.beta.core import clinical_events, patients

hba1c_codelist = codelist_from_csv("XXX", column="YYY")

dataset = create_dataset()
dataset.mean_hba1c = clinical_events.where(
        clinical_events.snomedct_code.is_in(hba1c_codelist)
).where(
        clinical_events.date.is_on_or_after("2022-07-01")
).numeric_value.mean_for_patient()
dataset.define_population(patients.exists_for_patient())
```

Performing arithmetic on numeric values of clinical

Finding the mean observed value of clinical events matching some crit

```
from ehrql import create_dataset, codelist_from_csv
from ehrql.tables.beta.core import clinical_events, patients

hba1c_codelist = codelist_from_csv("XXX", column="YYY")

dataset = create_dataset()
dataset.mean_hba1c = clinical_events.where(
        clinical_events.snomedct_code.is_in(hba1c_codelist)
).where(
        clinical_events.date.is_on_or_after("2022-07-01")
).numeric_value.mean_for_patient()
dataset.define_population(patients.exists_for_patient())
```

ehrQL **standardises** and makes EHR code:

- **readable**
- **reusable**
- **portable**

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable**

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **readable** for clinicians!

```
recent_asthma_rx_count = (
    prescriptions
    .where( prescriptions.dmd_code.is_in( asthma_medications ))
    .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
     .count_for_patient()
)
```

ehrQL is **reusable**

```
has_asthma = (
    asthma_diagnosis | ( recent_asthma_rx_count > 3 )
)
```

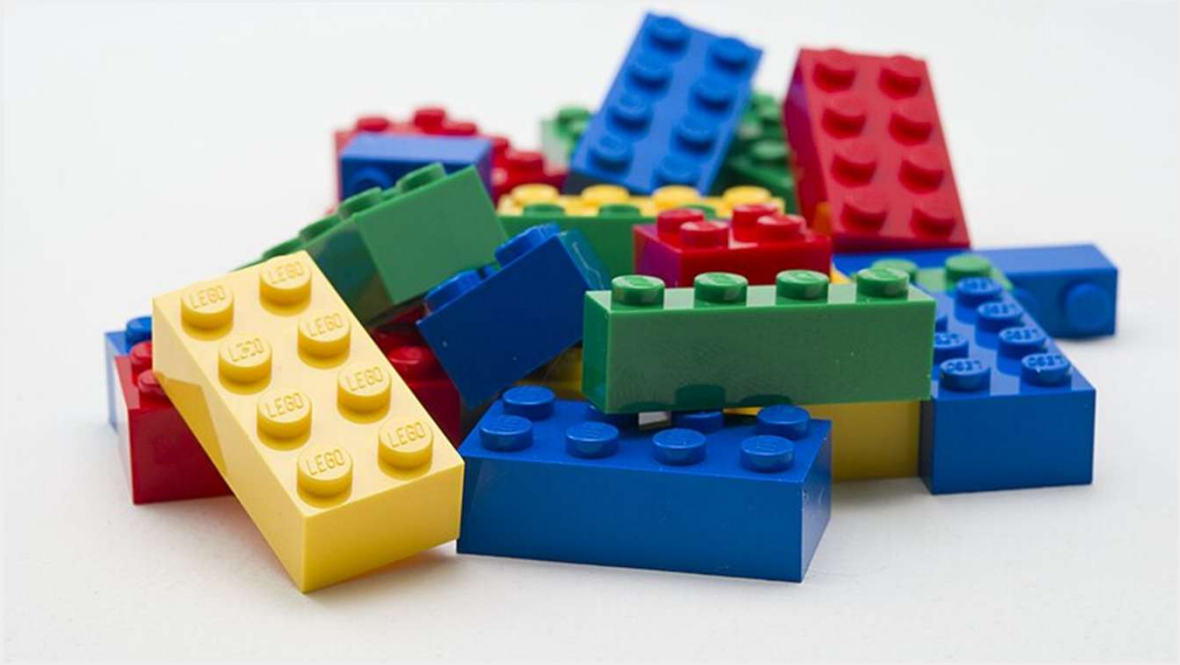
ehrQL is **reusable**

```
dataset.define_population(
    variables_library.has_asthma
)
```

ehrQL is **reusable**

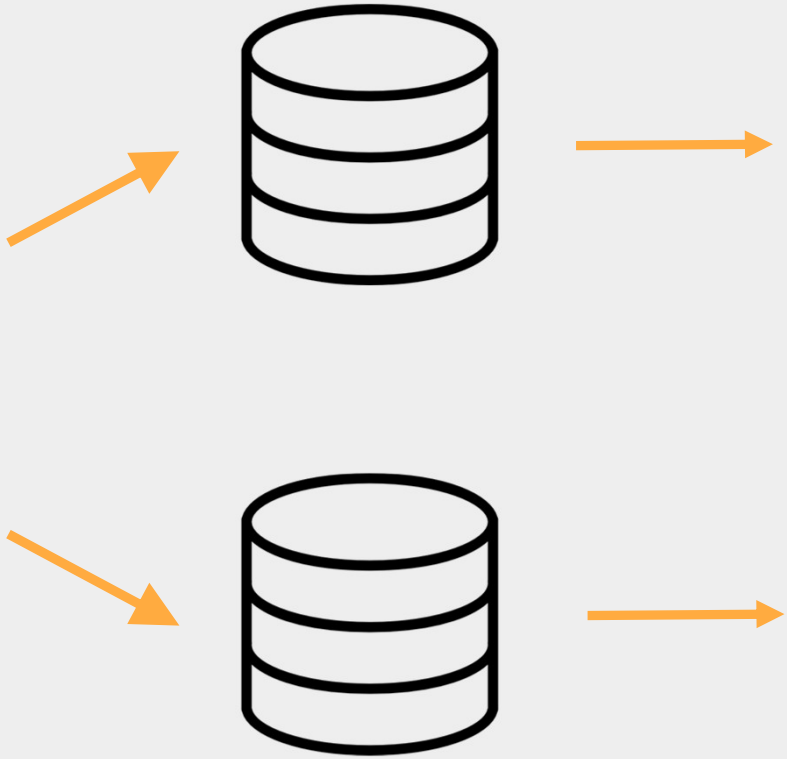
```
dataset.define_population(
    variables_library.has_recent_asthma(years=5)
)
```

ehrQL is **composable**



ehrQL is **portable**

```
# Diabetes mellitus variable library
# Business Rules for Quality and Outcomes Framework (QOF) 2021/22 (Version 46)
# Clinical data extraction criteria
from variable_lib_helper import (
    first_matching_event,
    last_matching_event,
    age_as_of,
    practice_registration_as_of,
    get_events_on_or_between,
)

import codelists

# Define a dataset including all variables needed for the diabetes mellitus (dm)
def make_dm_dataset(index_date):
    # Instantiate dataset
    dataset = create_dataset()

    # Extract prior events for further use in variable definitions below
    prior_events = clinical_events.where(
        clinical_events.date.is_on_or_before(index_date)
    )

    # Field number: 2
    # REG_DAT: The most recent date that the patient registered for GMS, where
    # this registration occurred on or before the achievement date.
    dataset.reg_dat = practice_registration_as_of(index_date).start_date

    # Field number: 3
    # REG_DAT: The first occurrence of the patient deregistering from GMS
    # following the latest GMS registration.
    dataset.dereg_dat = practice_registration_as_of(dataset.reg_dat).end_date

    # Field number: 4
    # PAT_AGE: The age of the patient in full years at the achievement date.
    dataset.pat_age = age_as_of(index_date)

    # Field number: 5
    # DM_DAT: Date of the first diabetes diagnosis up to and including the
    # achievement date.
    dataset.dm_dat = first_matching_event(prior_events, codelists.dm_cod).date

    # Field number: 6
    # DMLAT_DAT: Date of the most recent diabetes diagnosis up to and
    # including the achievement date.
    dataset.dmlat_dat = last_matching_event(prior_events, codelists.dm_cod).date

    # Field number: 7
    # DMRES_DAT: Date of the most recent diabetes diagnosis resolved code
    # recorded after the most recent diabetes diagnosis and up to and
    # including the achievement date.
    dataset.dmres_dat = last_matching_event(prior_events, codelists.dmres_cod).date

    # Field number: 8
    # DMMAX_DAT: Date of the most recent maximum tolerated diabetes treatment
    # code up to and including the achievement date.
    dataset.dmmax_dat = last_matching_event(prior_events, codelists.dmmax_cod).date

    # Field number: 9
    # IFCCHBA_DAT: Date of the most recent IFCC HbA1c monitoring range code up
```

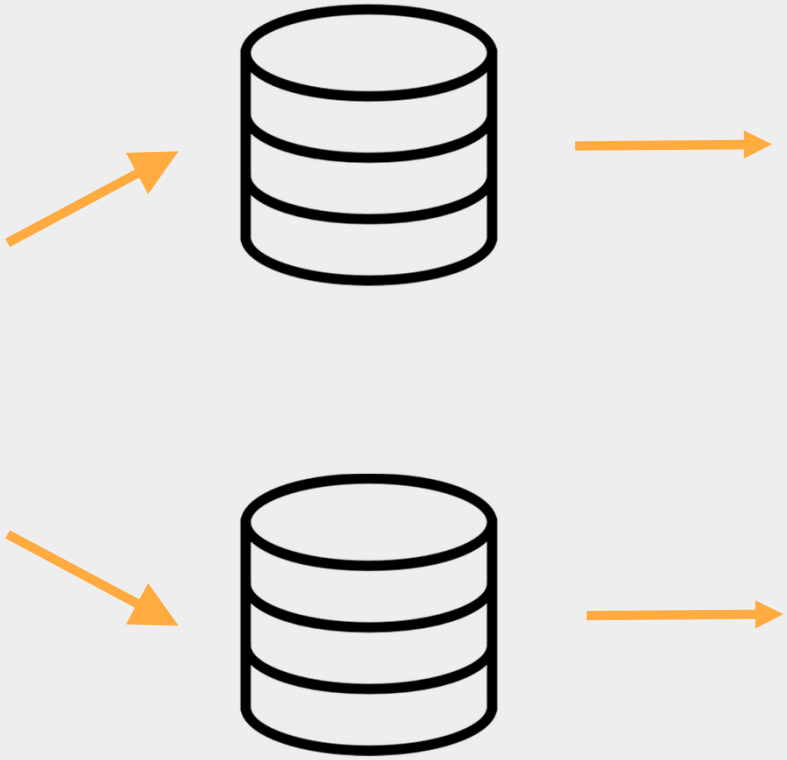


patient_id	has_asthma
375	NO
489	YES
611	NO

patient_id	has_asthma
204	YES
388	NO
525	NO

ehrQL is **portable**

```
recent_asthma_rx_count = (
  prescriptions
  .where( prescriptions.dmd_code.is_in( asthma_medications ))
  .where( prescriptions.date.is_after( "2024-11-25" - years(1) ))
  .count_for_patient()
)
```



patient_id	has_asthma
375	NO
489	YES
611	NO

patient_id	has_asthma
204	YES
388	NO
525	NO