Reac2Vac

Your Doctor's Guide to Vaccine Adverse Reactions

Fabien Plisson
Insight Health Data Science


Vaccines
save
2-3
MILLION
lives
each year
#VACCINESWORK
unicef
© UNICEF/AFGA2012-00020/Froutan

YET

...

Problem Statement:

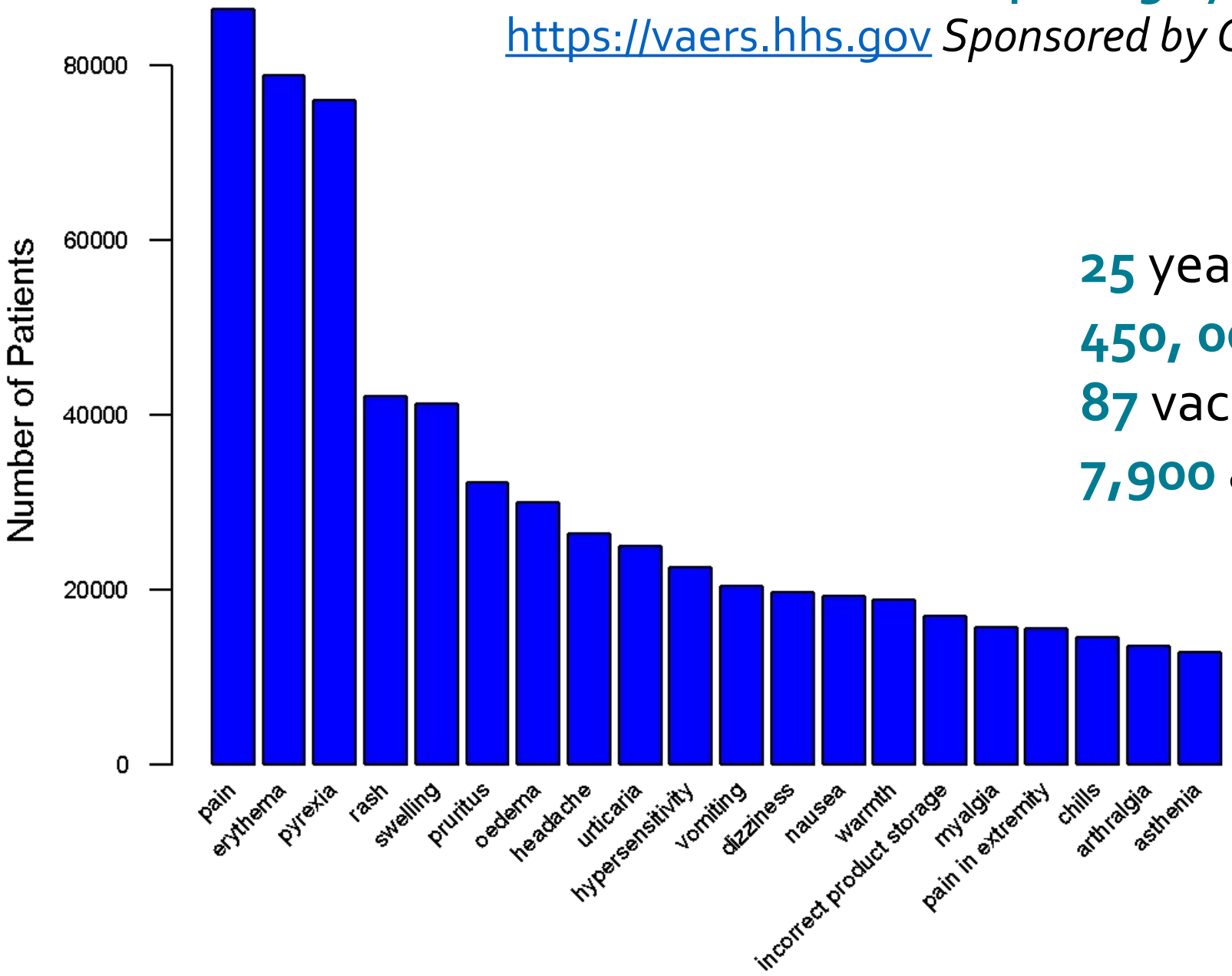
Can we predict the adverse events associated following immunization?

Target Audience:

Practitioners & Clinicians

Vaccine Adverse Event Reporting System (VAERS)

https://vaers.hhs.gov *Sponsored by CDC and FDA*



25 years in **75** Excel data sheets

450, 000 reports

87 vaccines type, n. repeats

7,900 adverse events

Patient Profile

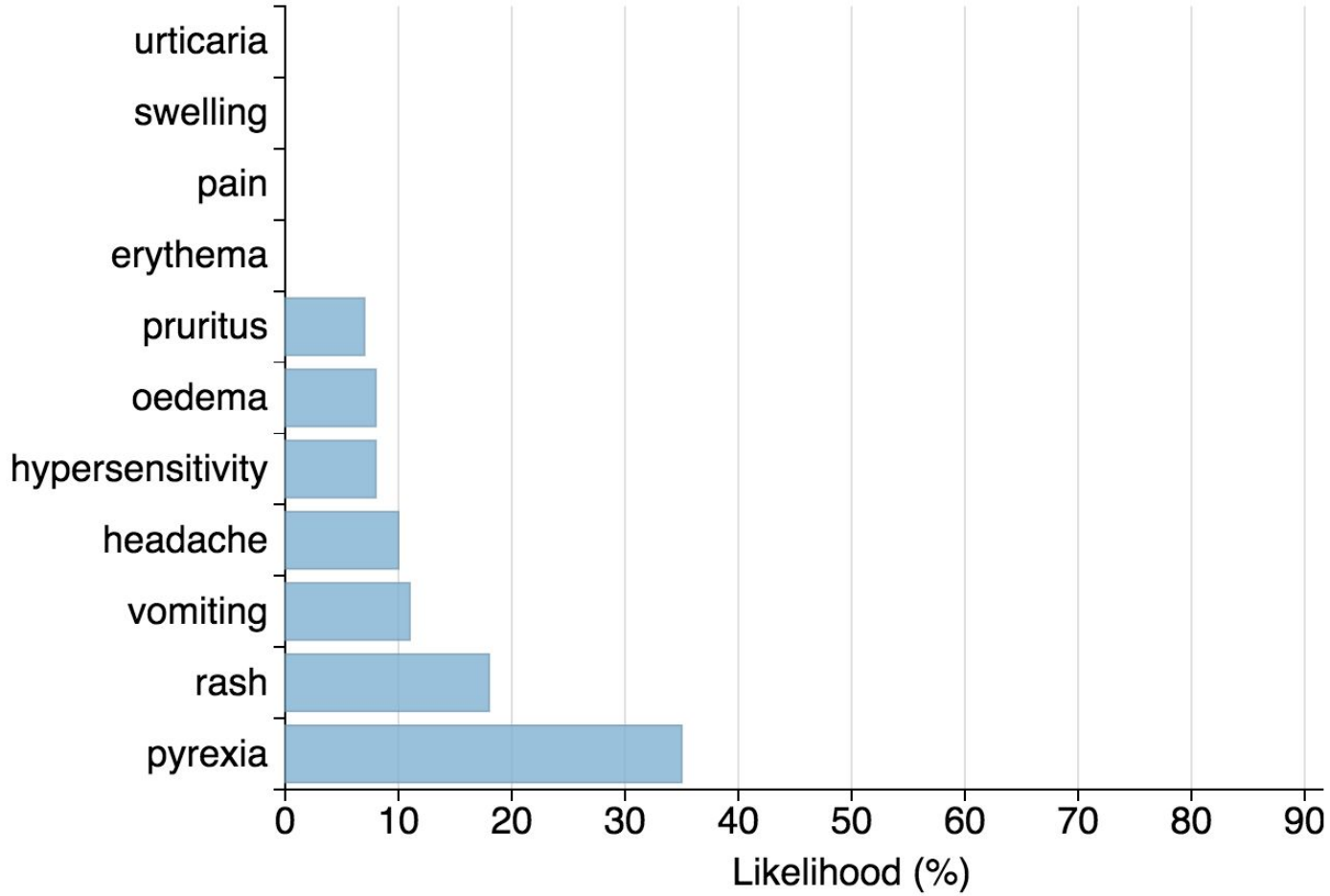
Medical Records Database
This test database has nearly 24,000 patient records, 9 patients profile are displayed below.

ID	Age	Sex	Emergency	Hospital (Days)	Death	Vaccine 1	Repeats 1	Vaccine 2	Repeats 2	Vaccine 3	Repeats 3
15658	72	F	No	0	No	Pneumonia	6	Diphtheria Tetanus Pertussis	6	Seasonal influenza	6
17898	16	F	No	0	No	Measles Mumps Rubella	7	Diphtheria Tetanus Pertussis	14	-	-
94564	0.3	M	Yes	1	No	Haemophilus influenza B	6	Diphtheria Tetanus Pertussis	13	-	-
4242	19	F	No	1	No	Hepatitis A/B	7	-	-	-	-
69572	40	F	No	1	No	Varicella	7	-	-	-	-
68532	24	M	Yes	1	No	Seasonal influenza	7	-	-	-	-
3838	4	M	No	0	No	Measles Mumps Rubella	7	Diphtheria Tetanus Pertussis	7	-	-
24494	66	M	No	1	No	Pneumonia	6	-	-	-	-
31021	56	F	No	14	No	Seasonal influenza	7	-	-	-	-

Enter your Patient VAERS ID: 68532 Run Predictions

68532

Predicted Adverse Events



WORKFLOW

Data Cleaning, Natural Language Processing, Regular Expression

Profile

id	state	age_yrs	died
25001	WI	0.2	0
25003	TX	0.8	1
25004	NY	0.9	0
25006	OH	16.7	0
25009	FL	3.3	0
25012	WI	0.2	0

+

Immunization

VAERS_ID	VAX_TYPE
25001	DTP
25003	DTP
25003	OPV
25004	OPV
25005	TD
25006	MMR

NLP

Features

```
u'age_yrs', u'died', u'er_visit', u'numdays', u'sex_f', u'sex_m',
u'sex_u', u'poliovirus_combos', u'haemophilus_influenza_b',
u'hepatitis', u'human_papillovirus', u'meningococcal_meningitis',
u'measles.mumps.rubella', u'pneumonia', u'rotavirus', u'typhoid',
u'anthrax', u'varicella', u'herpes', u'yellow_fever',
u'diphtheria.tetanus.pertussis', u'pandemic_influenza',
u'seasonal_influenza', u'hib_combos', u'pv_repeats', u'hib_repeats',
u'hep_repeats', u'hpv_repeats', u'mening_repeats', u'mmr_repeats',
u'pneum_repeats', u'rota_repeats', u'typh_repeats', u'anth_repeats',
u'vari_repeats', u'herp_repeats', u'yf_repeats', u'dtp_repeats',
u'paninfl_repeats', u'seasinfl_repeats', u'hib_combos_repeats'],
```

Adverse Events

VAERS_ID	SYMPTOMS
25001	Agitation
25003	Delirium, Hypokinesia, Hypotonia
25004	Chills, Dermatitis contact, Oedema genital, Pelvic pain
25005	Arthritis, Injection site oedema, Injection site reaction
25006	Convulsion, Dizziness
25007	Injection site inflammation, Injection site reaction

Clustering

Classes

id	SYMPTOMS	Class	PATCOUNT
25046	pain	2	86377
25074	pain	2	86377
25075	pain	2	86377
25075	pain	2	86377
25082	pain	2	86377
25093	pain	2	86377

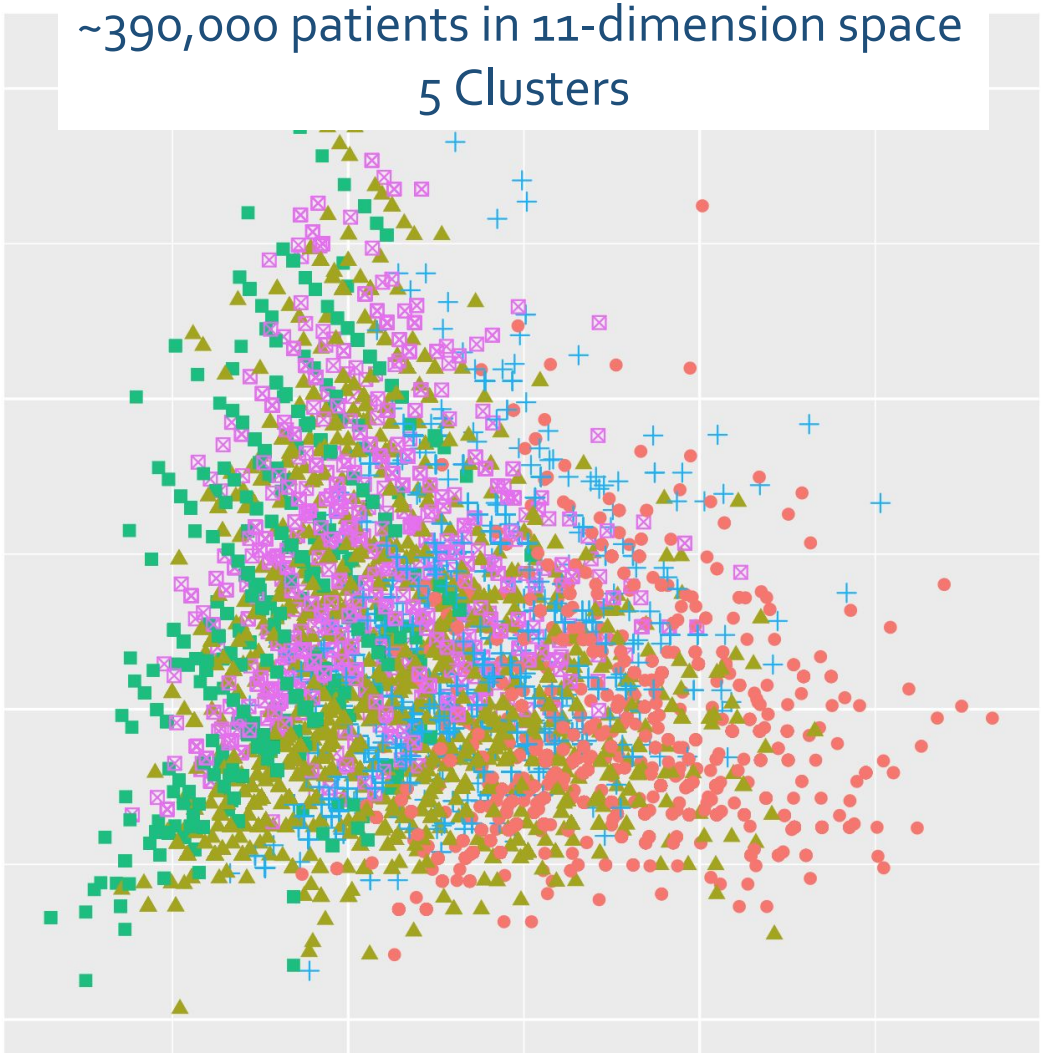
Likelihood Adverse Event

Return %
in each Class

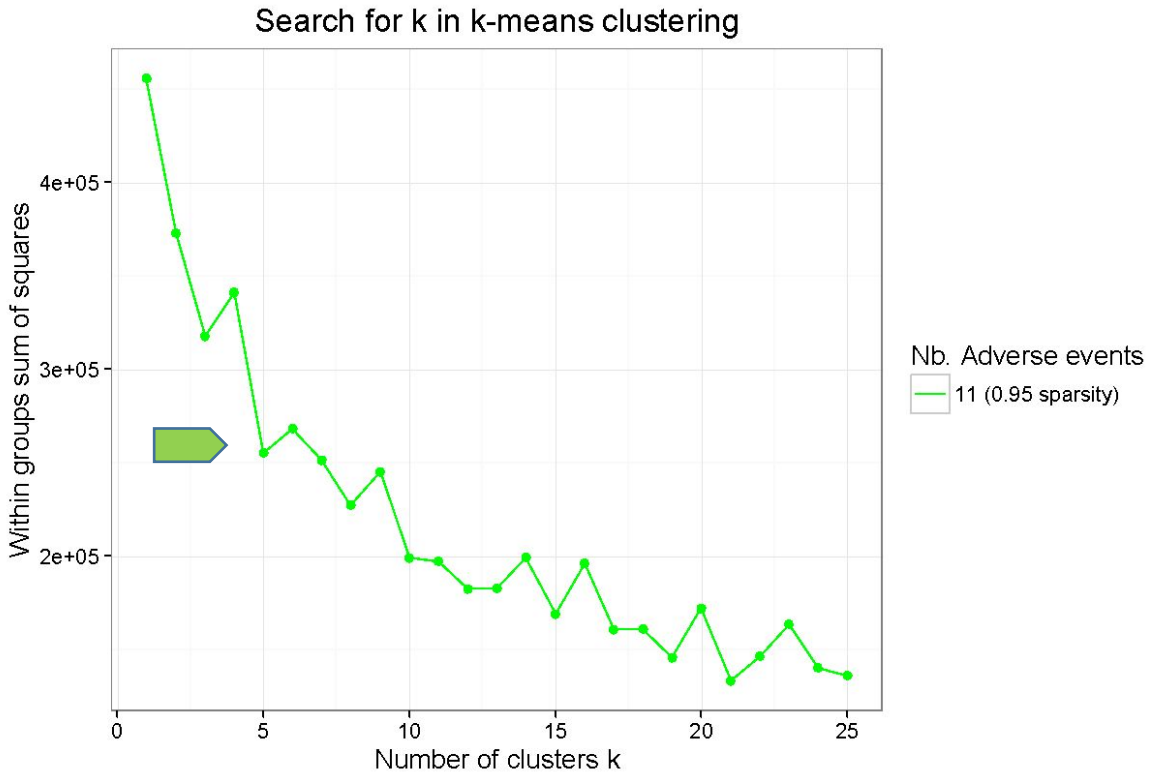
WORKFLOW

K-means Clustering Adverse Events

~390,000 patients in 11-dimension space
5 Clusters



Finding k: Elbow method



THE ALGORITHM

A Multiclass Solution

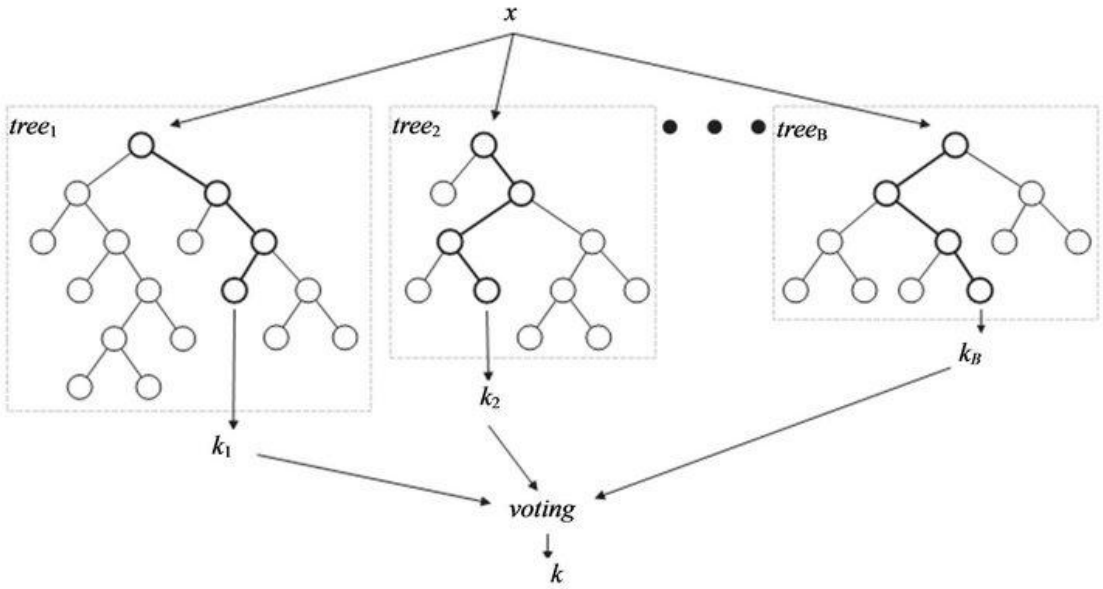
Random Forest

Why?

A multiclass classifier
Many Boolean values
Runs efficiently large datasets

How?

Imbalanced classes
Feature Selection (Recursive Feature Extraction)
Tuning: number of trees, depth
5-fold cross validation



Model	Test Accuracy (%)
Logistic Regression	29.4
K Nearest Neighbors	33.1
Naïve Bayes	21.7
Decision Tree (CART)	39.5
Random Forest	**40.1**

ABOUT ME

PhD Organic Chemistry
University of Queensland, Australia

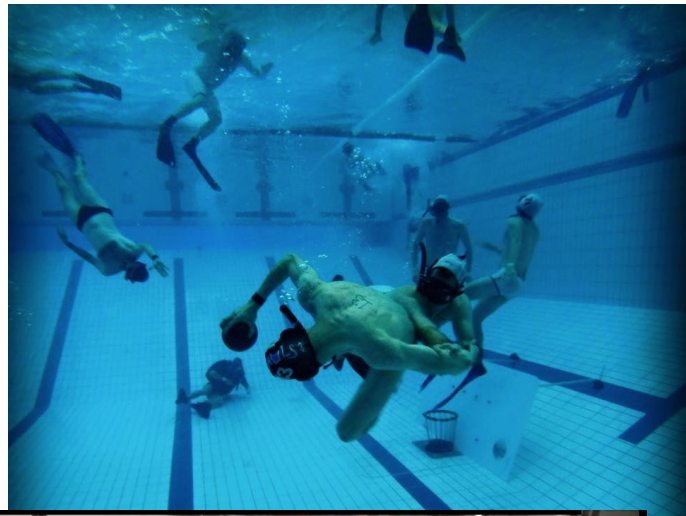
10 Years Research in Drug Discovery

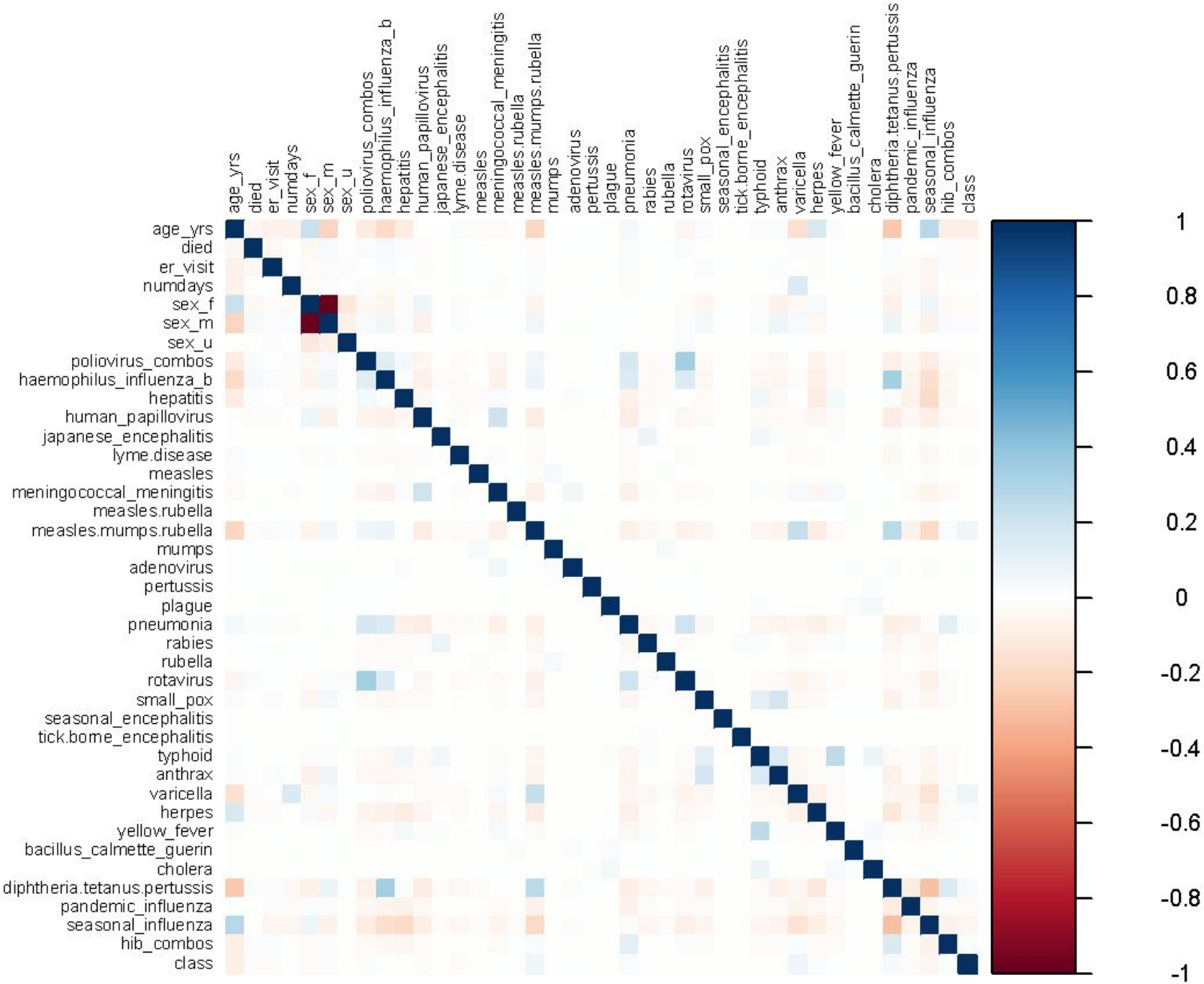
Pfizer
US

Protagonist Therapeutics
Australia

noscira
Spain

Pierre Fabre
France




age_yrs
died
er_visit
numdays
sex_f
sex_m
sex_u
poliovirus_combos
haemophilus_influenza_b
hepatitis
human_papillovirus
japanese_encephalitis
lyme.disease
measles
meningococcal_meningitis
measles.rubella
measles.mumps.rubella
mumps
adenovirus
pertussis
plague
pneumonia
rabies
rubella
rotavirus
small_pox
seasonal_encephalitis
tick.borne_encephalitis
typhoid
anthrax
varicella
herpes
yellow_fever
bacillus_calmette_guerin
cholera
diphtheria.tetanus.pertussis
pandemic_influenza
seasonal_influenza
hib_combos
class
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1