# Introduction to Artificial Intelligence and Machine Learning

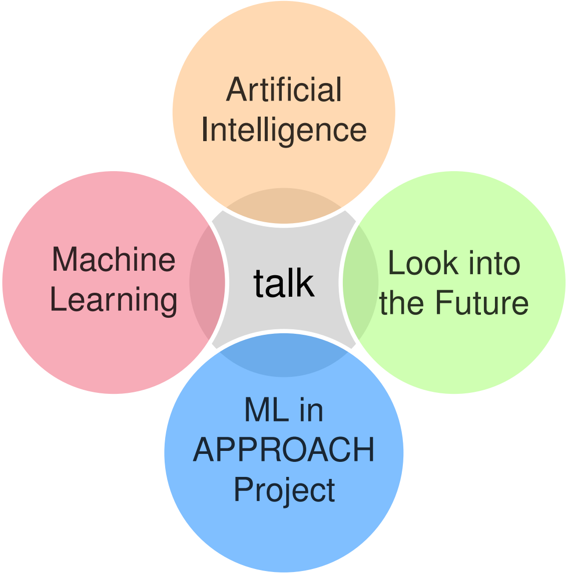
Paweł Widera

pawel.widera@ncl.ac.uk

Interdisciplinary Computing and Complex BioSystems group
School of Computing Science

**Lygature Partnerships MeetUp, Utrecht**
2019-10-29

# Outline

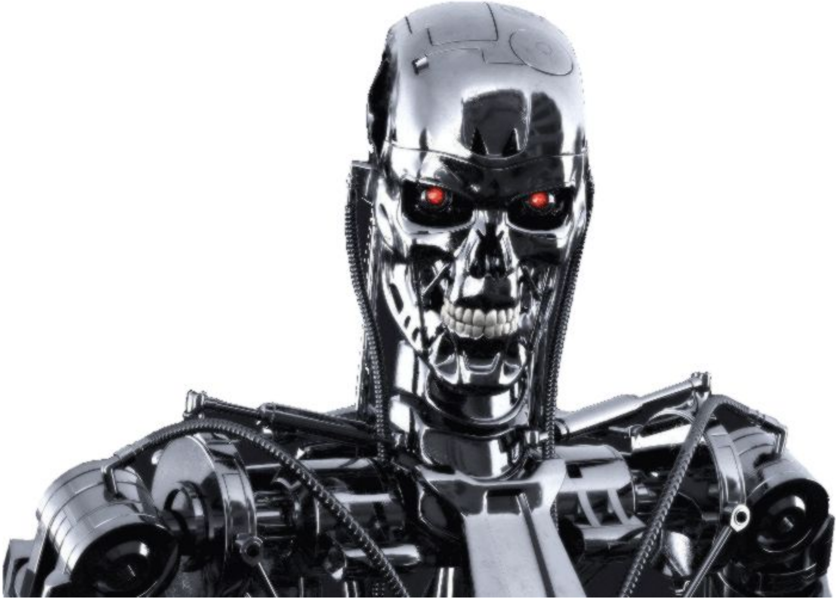


Artificial Intelligence
Machine Learning
talk
Look into the Future
ML in APPROACH Project

PART I

# Artificial Intelligence

# What is Artificial Intelligence?



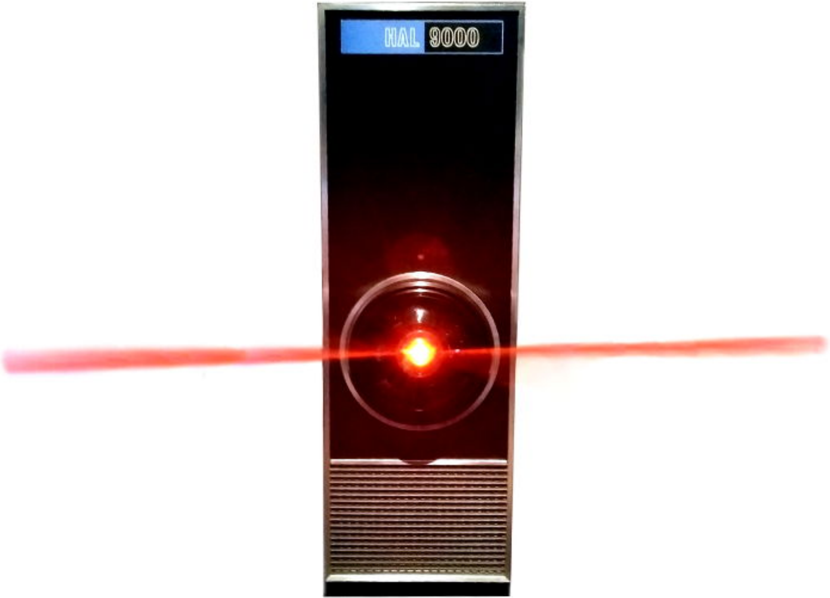
- T-800
- WALL-E
- HAL 9000
- smart speakers

# What is Artificial Intelligence?


- T-800
- WALL-E
- HAL 9000
- smart speakers

CC0 Lenin Estrada

# What is Artificial Intelligence?




- T-800
- WALL-E
- HAL 9000
- smart speakers

# What is Artificial Intelligence?


- T-800
- WALL-E
- HAL 9000
- smart speakers

(C) The Verge

# How fluid is the AI definition?

*"AI is whatever hasn't been done yet."*

(C) BBC / The Horizon Guide to AI

**Magic vs. computation**

- technological utopia of 1950s
- from robots to data
- from masterminds to assistants

# How fluid is the AI definition?

*“AI is whatever hasn’t been done yet.”*

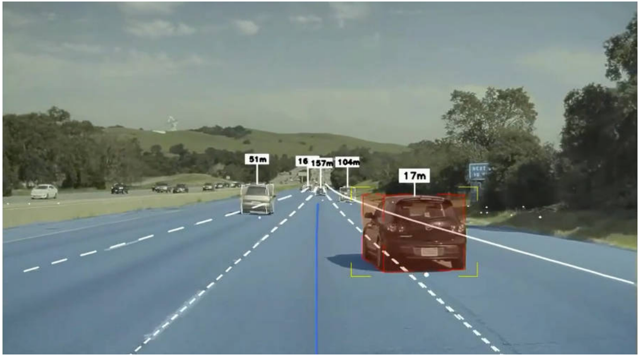
(C) Marek Lubas

**Magic vs. computation**

- technological utopia of 1950s
- from robots to data
- from masterminds to assistants

# How fluid is the AI definition?

*"AI is whatever hasn't been done yet."*




(C) Tesla / Autonomy Day

**Magic vs. computation**

- technological utopia of 1950s
- from robots to data
- from masterminds to assistants

# What is AI research about?

*"**Understanding data** — turning it into knowledge, conclusions, and actions."*

**Applications**

- puzzle solving
- planning
- moving
- recognizing objects
- speaking
- translating
- making art

# What is AI research about?

*“**Understanding data** — turning it into knowledge, conclusions, and actions.”*

## Applications

- puzzle solving
- planning
- moving
- recognizing objects
- speaking
- translating
- making art

# What is AI research about?

*“**Understanding data** — turning it into knowledge, conclusions, and actions.”*

style transfer [Gatys et al., 2015]

## Applications

- puzzle solving
- planning
- moving
- recognizing objects
- speaking
- translating
- making art

PART II

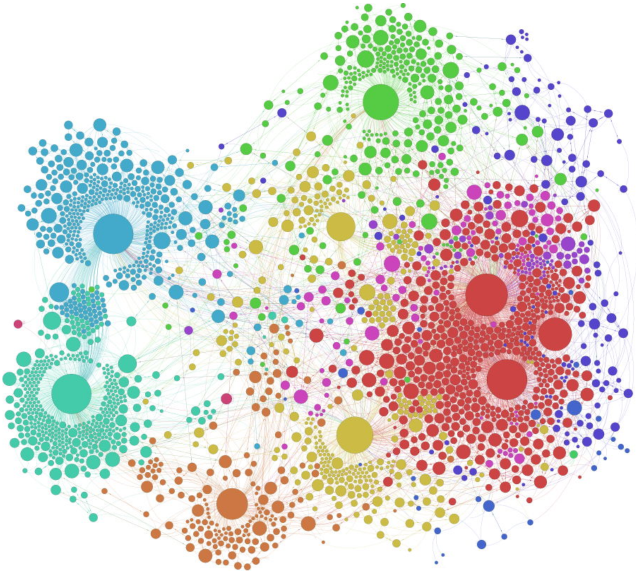
# Machine Learning

# What is Machine Learning?

*“ML is about making **AIs** automatically from the data,*
*so that they can improve with **experience**.”*

*“The ML algorithm is not a set of instructions,*
*it is the **learning process**.”*

# Main types of machine learning





## Unsupervised learning

network visualisation, finding patterns, making connections

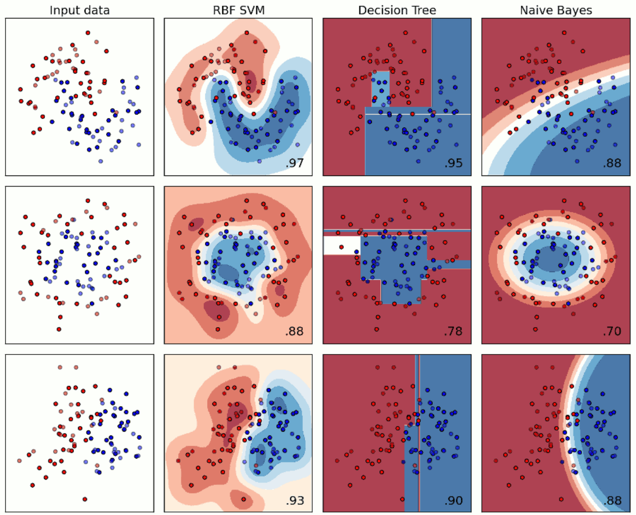
## Supervised learning

image classification, speech/text recognition, medical diagnosis, fraud detection, recommendation systems

## Reinforcement learning

traffic management, market analysis, robot navigation

# Main types of machine learning




generated with **scikit-learn**

## Unsupervised learning

network visualisation, finding patterns, making connections

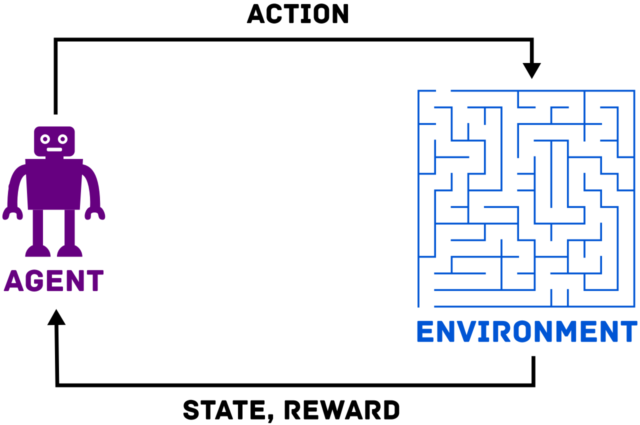
## Supervised learning

image classification, speech/text recognition, medical diagnosis, fraud detection, recommendation systems

## Reinforcement learning

traffic management, market analysis, robot navigation

# Main types of machine learning




## Unsupervised learning

network visualisation, finding patterns, making connections

## Supervised learning

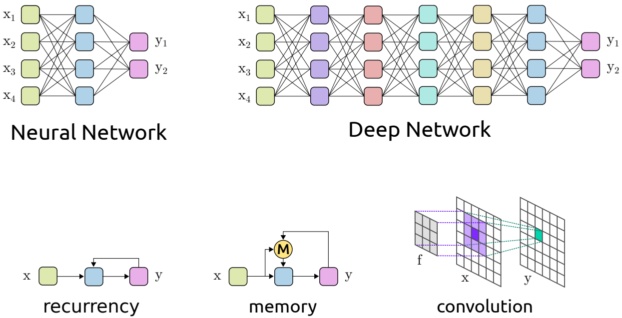
image classification, speech/text recognition, medical diagnosis, fraud detection, recommendation systems

## Reinforcement learning

traffic management, market analysis, robot navigation

# Deep learning

*Better name: "statistical inference from large datasets".*

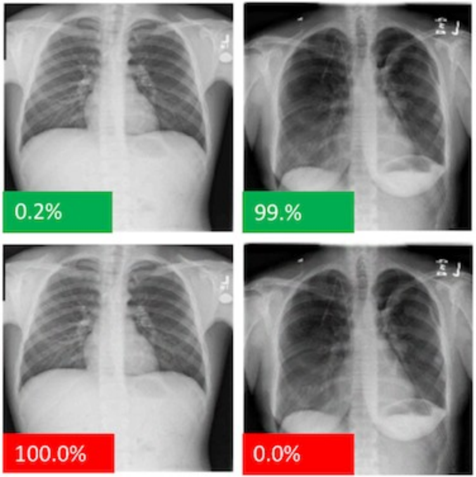



## In a nutshell

- pattern recognition engine applied to images, text or speech
- very effective for narrowly defined tasks
- prone to adversarial examples
- no explanation behind decisions (XAI)

# Deep learning

*Better name: “statistical inference from large datasets”.*



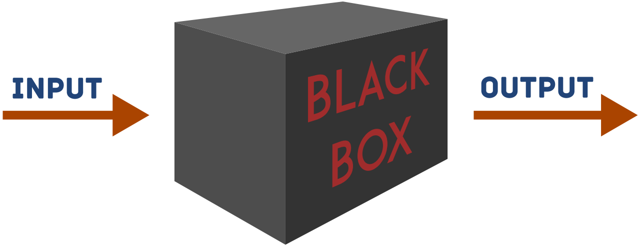

[Finlayson et al., 2019]

**In a nutshell**

- pattern recognition engine applied to images, text or speech
- very effective for narrowly defined tasks
- prone to adversarial examples
- no explanation behind decisions (XAI)

# Deep learning

*Better name: "statistical inference from large datasets".*




## In a nutshell

- pattern recognition engine applied to images, text or speech
- very effective for narrowly defined tasks
- prone to adversarial examples
- no explanation behind decisions (XAI)

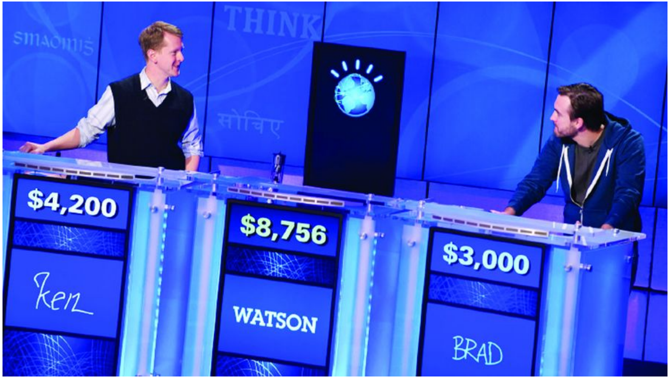
# Better than humans?


(C) Peter Morgan / REUTERS

## Events timeline

- 1997 — **chess**
  DeepBlue vs. Garry Kasparov
- 2011 — **Jeopardy!**
  Watson vs.
  Brad Rutter and Ken Jennings
- 2016 — **go**
  AlphaGo vs. Lee Sedol
- 2019 — **StarCraft**

# Better than humans?




(C) David Korchin

## Events timeline

- 1997 — **chess**
  DeepBlue vs. Garry Kasparov
- 2011 — **Jeopardy!**
  Watson vs.
  Brad Rutter and Ken Jennings
- 2016 — **go**
  AlphaGo vs. Lee Sedol
- 2019 — **StarCraft**

# Better than humans?


(C) Greg Kohs / AlphaGo documentary

## Events timeline

- 1997 — **chess**
  DeepBlue vs. Garry Kasparov
- 2011 — **Jeopardy!**
  Watson vs.
  Brad Rutter and Ken Jennings
- 2016 — **go**
  AlphaGo vs. Lee Sedol
- 2019 — **StarCraft**

# Better than humans?

(C) Greg Kohs / AlphaGo documentary

## Events timeline

- 1997 — **chess**
  DeepBlue vs. Garry Kasparov
- 2011 — **Jeopardy!**
  Watson vs.
  Brad Rutter and Ken Jennings
- 2016 — **go**
  AlphaGo vs. Lee Sedol
- 2019 — **StarCraft**

AlphaGo does $\approx$ 50 moves in 4h game. We make 50 decisions per second using higher level of abstraction.

# Better than humans?

(C) Blizzard Entertainment

## Events timeline

- 1997 — **chess**
  DeepBlue vs. Garry Kasparov
- 2011 — **Jeopardy!**
  Watson vs.
  Brad Rutter and Ken Jennings
- 2016 — **go**
  AlphaGo vs. Lee Sedol
- 2019 — **StarCraft**

AlphaGo does $\approx$ 50 moves in 4h game. We make 50 decisions per second using higher level of abstraction.

# What can we expect from ML in medicine?

(C) Paramount Television / Star Trek Voyager

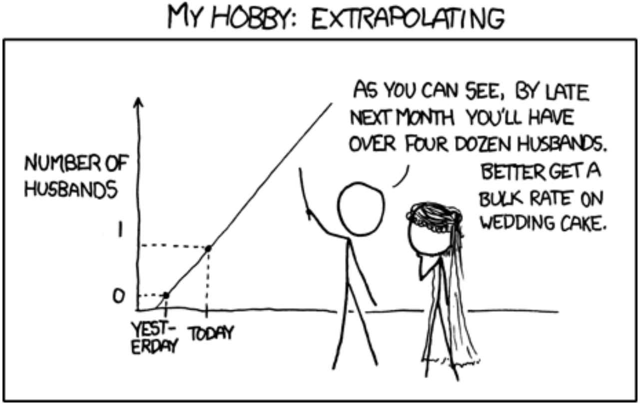
### Promise of science-fiction

- Star Trek's tricorder
- danger of extrapolation

### Failure of IBM Watson

- applied to oncology
- read and aggregated the corpus of medical literature (10 000 human diseases)
- problem — Watson cannot **read** unstructured text
- small success: matched patients missed for clinical trial

# What can we expect from ML in medicine?




https://xkcd.com/605/

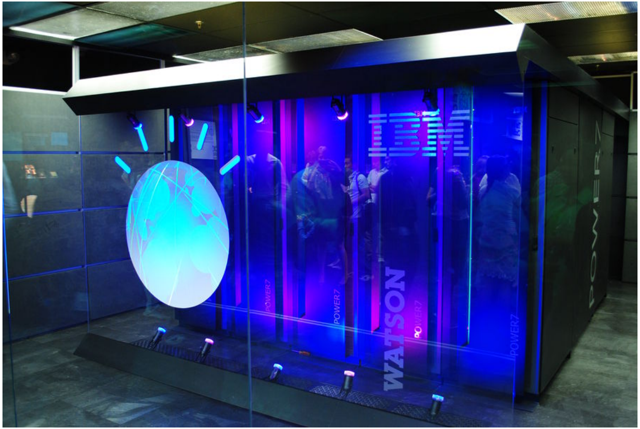
**Promise of science-fiction**

- Star Trek's tricorder
- danger of extrapolation

**Failure of IBM Watson**

- applied to oncology
- read and aggregated the corpus of medical literature (10 000 human diseases)
- problem — Watson cannot **read** unstructured text
- small success: matched patients missed for clinical trial

# What can we expect from ML in medicine?






https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care

## Promise of science-fiction

- Star Trek's tricorder
- danger of extrapolation

## Failure of IBM Watson

- applied to oncology
- read and aggregated the corpus of medical literature (10 000 human diseases)
- problem — Watson cannot **read** unstructured text
- small success: matched patients missed for clinical trial

# Reading comprehension benchmark

*"The large **ball** crashed right through the **table** because **it** was made of styrofoam."*

***Does "it" refer to "ball"?***

---

*"I put the **cake** away in the **refrigerator**. **It** has a lot of butter in it."*

***Does "it" refer to "cake"?***

**SuperGLUE** **[Wang et al., 2019]**

- example questions
- leaderboard
- dataset selection

# Reading comprehension benchmark

*"The large **ball** crashed right through the **table** because **it** was made of styrofoam."*

***Does "it" refer to "ball"?***

---

*"I put the **cake** away in the **refrigerator**. **It** has a lot of butter in it."*

***Does "it" refer to "cake"?***

**SuperGLUE [Wang et al., 2019]**

- example questions
- leaderboard
- dataset selection

# Reading comprehension benchmark

| rank | name | algorithm | score |
|---|---|---|---|
| 1 | human baseline | — | 89.8 |
| 2 | Google Brain | T5 | 88.9 |
| 3 | Facebook AI | RoBERTa | 84.6 |
| 4 | IBM Research AI | BERT-mtl | 73.5 |
| 5 | baseline | BERT++ | 71.5 |

`https://super.gluebenchmark.com/leaderboard`

**SuperGLUE** **[Wang et al., 2019]**

- example questions
- leaderboard
- dataset selection

# Reading comprehension benchmark

| rank | name | algorithm | score |
|---|---|---|---|
| 1 | human baseline | — | 89.8 |
| 2 | Google Brain | T5 | 88.9 |
| 3 | Facebook AI | RoBERTa | 84.6 |
| 4 | IBM Research AI | BERT-mtl | 73.5 |
| 5 | baseline | BERT++ | 71.5 |

https://super.gluebenchmark.com/leaderboard

**SuperGLUE [Wang et al., 2019]**

- example questions
- leaderboard
- dataset selection

"We turned down many medical text datasets because they are usually only accessible with explicit permission and credentials from the data owners."

# What is ML good at today?

## Computer vision / image analysis

- body scans, retina images, skin lesions, pathology slides
- reduces false omissions from 30% to single digits
- better and faster than human experts, will become a standard

## Analysis of high throughput data

- genomics, proteomics, lipidomics, metabolomics
- polygenic risk score (genome variations associated to diseases)

## Decision support

- e.g. selection of patients for clinical trials

# What is ML good at today?

## Computer vision / image analysis

- body scans, retina images, skin lesions, pathology slides
- reduces false omissions from 30% to single digits
- better and faster than human experts, will become a standard

## Analysis of high throughput data

- genomics, proteomics, lipidomics, metabolomics
- polygenic risk score (genome variations associated to diseases)

## Decision support

- e.g. selection of patients for clinical trials

# What is ML good at today?

## Computer vision / image analysis

- body scans, retina images, skin lesions, pathology slides
- reduces false omissions from 30% to single digits
- better and faster than human experts, will become a standard

## Analysis of high throughput data

- genomics, proteomics, lipidomics, metabolomics
- polygenic risk score (genome variations associated to diseases)

## Decision support

- e.g. selection of patients for clinical trials

PART III

# ML in APPROACH project

## Acknowledgement

The research leading to these results has received support from the **Innovative Medicines Initiative Joint Undertaking** under grant **Agreement no. 115770**, resources of which are composed of financial contribution from the **European Union's Seventh Framework Programme** (FP7/2007-2013) and **EFPIA** companies' in kind contribution. See: https://www.imi.europa.eu.

## Disclaimer

This communication reflects the views of the **APPROACH consortium** and neither **IMI** nor the **European Union** and **EFPIA** are liable for any use that may be made of the information contained herein.

## WP1 Leaders

Christoph Ladel
John Loughlin

## WP3 Leaders

Floris Lafeber
Agnès Lalande

## Data analysis / Machine learning

Jaume Bacardit
Paweł Widera
Paco Welsing

## Data harmonisation

Samuel Danso
Paweł Widera
Paco Welsing
Sjaak Peelen

## CHECK

Anne Marijnissen
Eefje van Helvoort

## MUST

Ida Haugen
Janicke Magnus
Mari Skinnes

## DIGICOD

Francis Berenbaum
Jeremie Sellam

## PROCOAC

Francisco Blanco
Joana Silva
Carlos Tilve

## HOSTAS

Margreet Kloppenburg
Marieke Leof
Jose Krol

## Principal Investigators

Jonathan Larkin
Harrie Weinans

# APPROACH objectives

Recruit → Observe → Discover

- recruit **patients** for a new clinical study
- observe the **disease progression** (collecting data)
- discover **mechanisms** behind progression

## Long-term goals

1. better **definition** of the disease and its subtypes
2. more effective, personalised **treatment**

# APPROACH objectives

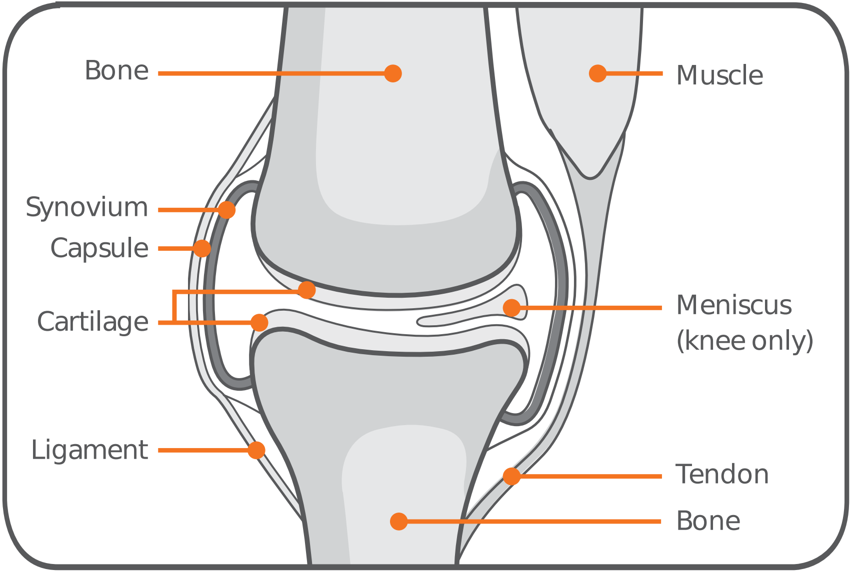
Recruit → Observe → Discover

- recruit **patients** for a new clinical study
- observe the **disease progression** (collecting data)
- discover **mechanisms** behind progression

## Long-term goals

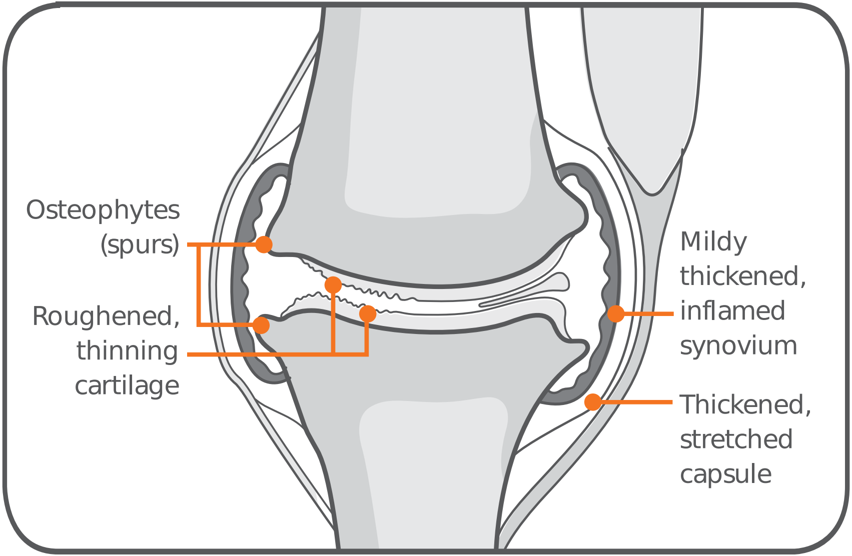
1. better **definition** of the disease and its subtypes
2. more effective, personalised **treatment**

# Osteoarthritis (OA)



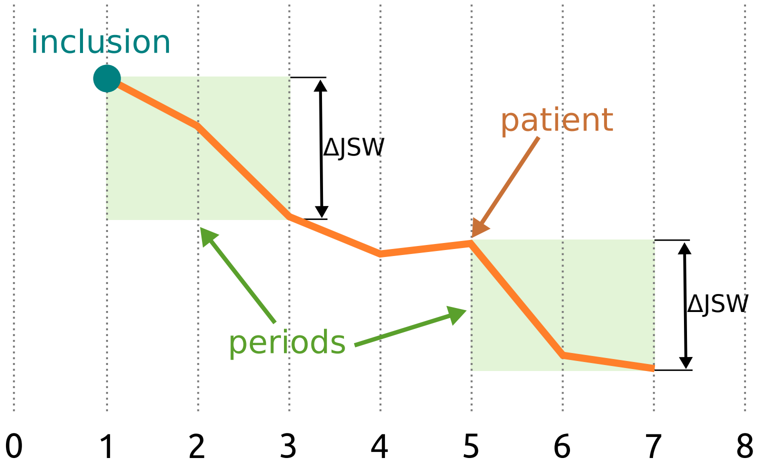

https://www.arthritisresearchuk.org/

# Osteoarthritis (OA)




https://www.arthritisresearchuk.org/

# Patient data



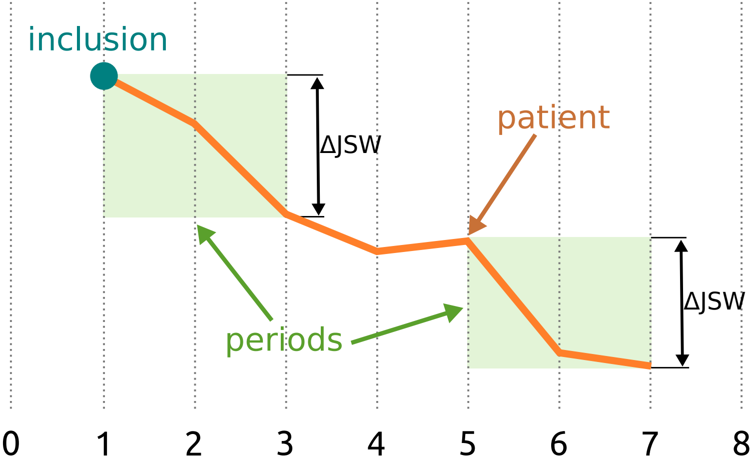

## Periods

- duration: **2+ years**
- OA **clinical criteria** satisfied at inclusion

## Categories

- **N**: non-progressors
- **P**: increased/stable pain
- **S**: structural progression
- **P+S**: both

# Patient data




## Periods

- duration: **2+ years**
- OA **clinical criteria** satisfied at inclusion

## Categories

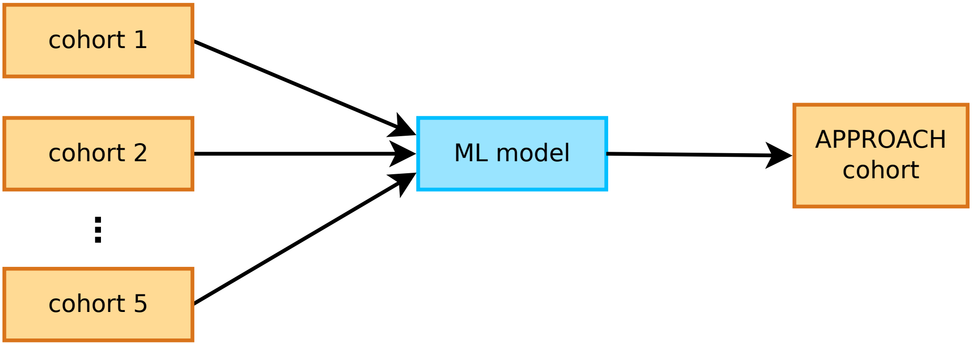
- **N**: non-progressors
- **P**: increased/stable pain
- **S**: structural progression
- **P+S**: both

# APPROACH cohort

- **new** longitudinal cohort focused on **fast** knee OA progressors
- recruited from **5 existing** European OA cohorts using **machine learning** models trained to estimate **progression probability**

| cohort | location | patients | attributes | visits | activity |
|---|---|---:|---:|---:|---:|
| MUST | Oslo, NO | 630 | 886 | 1 | 2010–2013 |
| HOSTAS | Leiden, NL | 538 | 130 | 3 | 2009–2015 |
| DIGICOD | Paris, FR | 377 | 425 | 2 | 2013–2016 |
| PROCOAC | A Coruña, ES | 983 | 288 | 7 | 2002–2016 |
| CHECK | Utrecht, NL | 1002 | 512 | 10 | 2006–2016 |

# Recruitment idea



# Challenges arising from the data

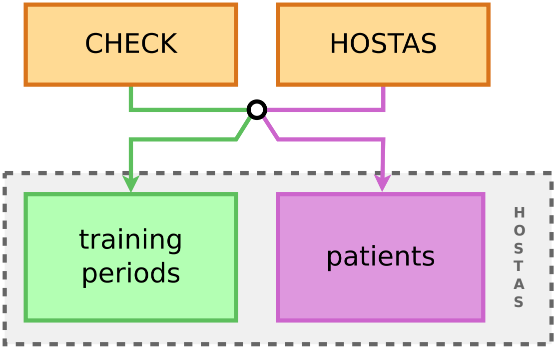
differences in collected data

differences in visits timing

decisions on out-dated information

1. cohorts other than CHECK **cannot** be used in **model training** (not enough data to define progression)
2. the **date of last visit** differs across cohorts and patients
3. model **prediction confidence** gets worse for **older** data

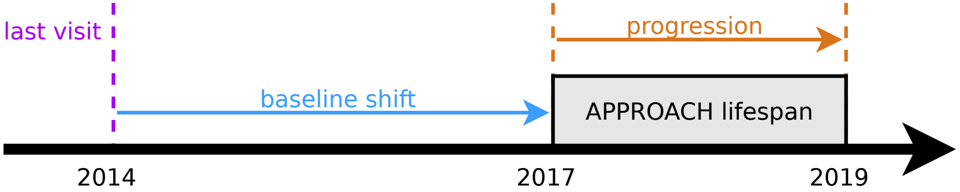
# Harmonisation to CHECK




**Harmonisation**

- **semantic** processing
  - common **data concepts**
  - **mapping** between the concepts
- **syntactic** processing
  - data **transformation**

# Baseline shift



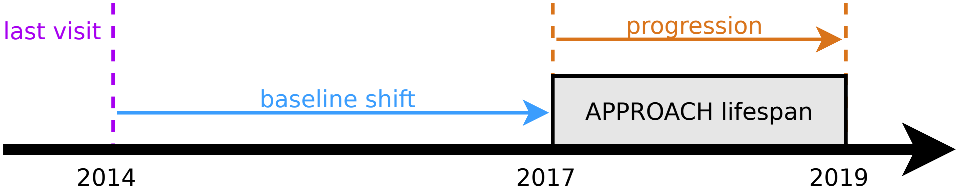

## Model training

- using harmonised CHECK periods
- baseline shift of 0, 2, 3 or 5 years

## Consequences

- need to train **separate** model for each **cohort** and each **shift**

# Baseline shift



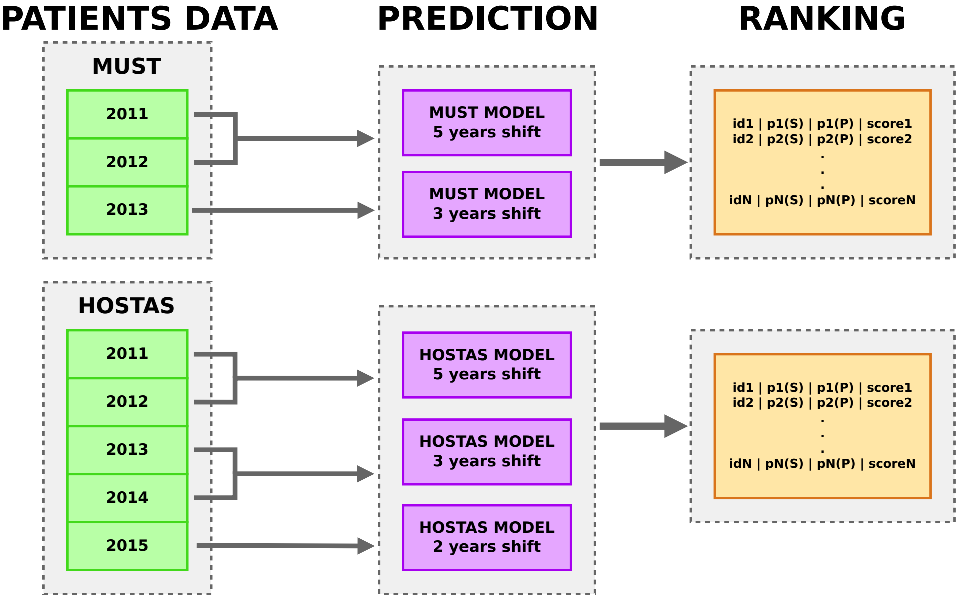

## Model training

- using harmonised CHECK periods
- baseline shift of 0, 2, 3 or 5 years

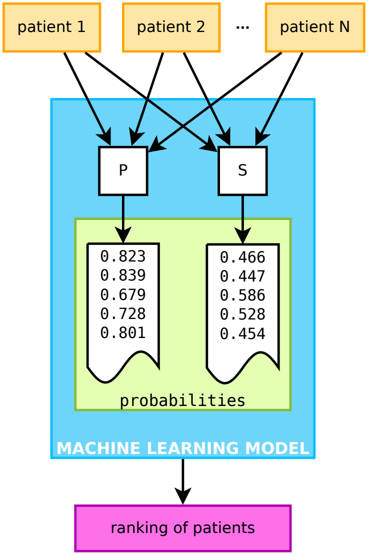
## Consequences

- need to train **separate** model for each **cohort** and each **shift**

# Multi-model prediction



# Ranking based on model confidence




- the ML model is composed of two **sub-models**, separately predicting **pain** (P) and **structure** (S) related progression
- the **outcome** of the model is per patient **probability** of becoming an OA progressor
- the probabilities are used to **rank** the patients
- **top patients** in the ranking are **more likely** to progress

# Example of combined ranking

**Top 16**

| ID | P | S | year | score |
|---|---|---|---|---|
| 249 | 0.703 | 0.238 | 2013 | 0.617390 |
| 389 | 0.583 | 0.226 | 2012 | 0.477706 |
| 626 | 0.493 | 0.232 | 2013 | 0.475673 |
| 370 | 0.588 | 0.242 | 2011 | 0.441096 |
| 461 | 0.616 | 0.122 | 2012 | 0.435782 |
| 450 | 0.539 | 0.182 | 2012 | 0.425743 |
| 597 | 0.529 | 0.185 | 2012 | 0.421610 |
| 546 | 0.532 | 0.181 | 2012 | 0.421019 |
| 532 | 0.565 | 0.145 | 2012 | 0.419248 |
| 92 | 0.555 | 0.224 | 2011 | 0.413993 |
| 376 | 0.499 | 0.192 | 2012 | 0.408029 |
| 385 | 0.530 | 0.158 | 2012 | 0.406257 |
| 397 | 0.510 | 0.178 | 2012 | 0.406257 |
| 505 | 0.470 | 0.214 | 2012 | 0.403895 |
| 484 | 0.520 | 0.157 | 2012 | 0.399762 |
| 177 | 0.436 | 0.240 | 2012 | 0.399171 |

**Bottom 16**

| ID | P | S | year | score |
|---|---|---|---|---|
| 78 | 0.062 | 0.221 | 2011 | 0.150398 |
| 241 | 0.074 | 0.209 | 2011 | 0.150398 |
| 583 | 0.061 | 0.193 | 2012 | 0.149984 |
| 447 | 0.076 | 0.177 | 2012 | 0.149394 |
| 297 | 0.110 | 0.168 | 2011 | 0.147741 |
| 242 | 0.050 | 0.227 | 2011 | 0.147209 |
| 551 | 0.050 | 0.197 | 2012 | 0.145851 |
| 268 | 0.046 | 0.226 | 2011 | 0.144552 |
| 368 | 0.082 | 0.159 | 2012 | 0.142308 |
| 250 | 0.058 | 0.204 | 2011 | 0.139238 |
| 569 | 0.079 | 0.152 | 2012 | 0.136403 |
| 599 | 0.077 | 0.130 | 2013 | 0.135813 |
| 307 | 0.084 | 0.166 | 2011 | 0.132860 |
| 69 | 0.047 | 0.197 | 2011 | 0.129672 |
| 548 | 0.065 | 0.178 | 2011 | 0.129140 |
| 302 | 0.072 | 0.143 | 2011 | 0.114260 |

**MUST**

| | |
|---|---|
| 2011 | 260 |
| 2012 | 254 |
| 2013 | 25 |

- when **combining rankings**, models using **older** data are trusted **less**

# Example of combined ranking

### Top 16

| ID | P | S | year | score |
|---|---|---|---|---|
| 249 | 0.703 | 0.238 | 2013 | 0.617390 |
| 389 | 0.583 | 0.226 | 2012 | 0.477706 |
| 626 | 0.493 | 0.232 | 2013 | 0.475673 |
| 370 | 0.588 | 0.242 | 2011 | 0.441096 |
| 461 | 0.616 | 0.122 | 2012 | 0.435782 |
| 450 | 0.539 | 0.182 | 2012 | 0.425743 |
| 597 | 0.529 | 0.185 | 2012 | 0.421610 |
| 546 | 0.532 | 0.181 | 2012 | 0.421019 |
| 532 | 0.565 | 0.145 | 2012 | 0.419248 |
| 92 | 0.555 | 0.224 | 2011 | 0.413993 |
| 376 | 0.499 | 0.192 | 2012 | 0.408029 |
| 385 | 0.530 | 0.158 | 2012 | 0.406257 |
| 397 | 0.510 | 0.178 | 2012 | 0.406257 |
| 505 | 0.470 | 0.214 | 2012 | 0.403895 |
| 484 | 0.520 | 0.157 | 2012 | 0.399762 |
| 177 | 0.436 | 0.240 | 2012 | 0.399171 |

### Bottom 16

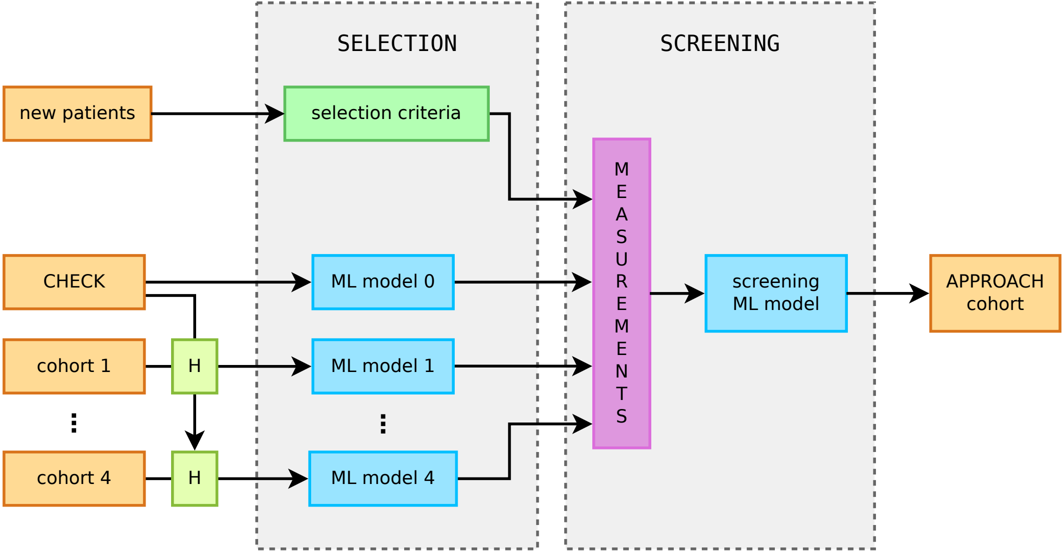
| ID | P | S | year | score |
|---|---|---|---|---|
| 78 | 0.062 | 0.221 | 2011 | 0.150398 |
| 241 | 0.074 | 0.209 | 2011 | 0.150398 |
| 583 | 0.061 | 0.193 | 2012 | 0.149984 |
| 447 | 0.076 | 0.177 | 2012 | 0.149394 |
| 297 | 0.110 | 0.168 | 2011 | 0.147741 |
| 242 | 0.050 | 0.227 | 2011 | 0.147209 |
| 551 | 0.050 | 0.197 | 2012 | 0.145851 |
| 268 | 0.046 | 0.226 | 2011 | 0.144552 |
| 368 | 0.082 | 0.159 | 2012 | 0.142308 |
| 250 | 0.058 | 0.204 | 2011 | 0.139238 |
| 569 | 0.079 | 0.152 | 2012 | 0.136403 |
| 599 | 0.077 | 0.130 | 2013 | 0.135813 |
| 307 | 0.084 | 0.166 | 2011 | 0.132860 |
| 69 | 0.047 | 0.197 | 2011 | 0.129672 |
| 548 | 0.065 | 0.178 | 2011 | 0.129140 |
| 302 | 0.072 | 0.143 | 2011 | 0.114260 |

**MUST**

| | |
|---|---|
| 2011 | 260 |
| 2012 | 254 |
| 2013 | 25 |

- when **combining rankings**, models using **older** data are trusted **less**

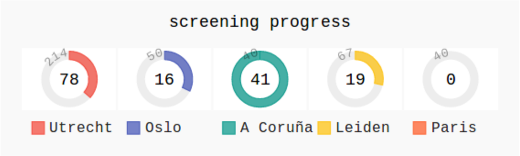
# Two-stage recruitment process

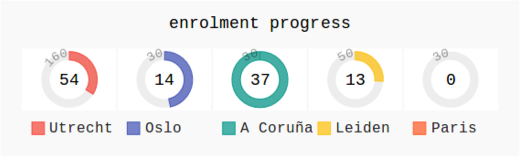


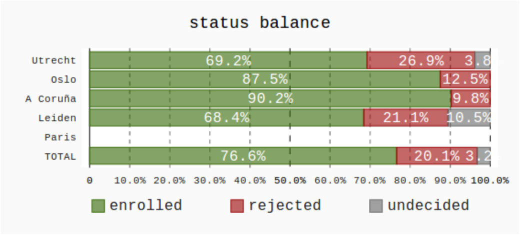
# Decision support and progress monitoring

| ID | RC | screening | P | S | score | decision | choice |
|---|---|---|---|---|---|---|---|
| 0015280201 | U | 44d (28 Feb) | 0.131 | 0.403 | -1.000 | 2018-03-15 | enrol |
| 0015281401 | L | 18d (26 Mar) | 0.150 | 0.395 | -1.013 | | ✓ |
| 0015280201 | U | 16d (28 Mar) | 0.196 | 0.378 | -1.015 | 2018-04-05 | enrol |
| 0015280201 | U | 56d (16 Feb) | 0.155 | 0.391 | -1.040 | 2018-03-15 | enrol |
| 0015280201 | U | 28d (16 Mar) | 0.228 | 0.361 | -1.081 | | ✓ |
| 0015280201 | U | 70d (02 Feb) | 0.248 | 0.353 | -1.090 | 2018-03-15 | enrol |
| 0015280201 | U | 63d (09 Feb) | 0.144 | 0.389 | -1.114 | 2018-03-15 | enrol |
| 0015280201 | U | 7d (06 Apr) | 0.223 | 0.358 | -1.141 | | |
| 0015280201 | U | 70d (02 Feb) | 0.206 | 0.361 | -1.181 | 2018-03-22 | reject |
| 0015280201 | U | 63d (09 Feb) | 0.205 | 0.359 | -1.210 | 2018-03-29 | reject |
| 0015280201 | U | 16d (28 Mar) | 0.543 | 0.233 | -1.243 | | |
| 0015280201 | U | 70d (02 Feb) | 0.188 | 0.362 | -1.250 | 2018-03-22 | reject |
| 0015280201 | U | 23d (21 Mar) | 0.372 | 0.294 | -1.260 | | |
| 0015280201 | U | 84d (19 Jan) | 0.172 | 0.367 | -1.260 | 2018-03-01 | reject |
| 0015280201 | U | 70d (02 Feb) | 0.151 | 0.373 | -1.281 | 2018-03-22 | reject |
| 0015280201 | U | 44d (28 Feb) | 0.197 | 0.354 | -1.308 | 2018-04-12 | reject |
| 0015781901 | O | 52d (20 Feb) | 0.271 | 0.327 | -1.308 | 2018-04-05 | reject |
| 0015280201 | U | 77d (26 Jan) | 0.146 | 0.369 | -1.353 | 2018-03-15 | reject |
| 0017241701 | A | 50d (22 Feb) | 0.159 | 0.364 | -1.356 | 2018-04-04 | reject |
| 0015280201 | U | 7d (06 Apr) | 0.434 | 0.262 | -1.377 | | ✗ |
| 0015280201 | U | 63d (09 Feb) | 0.175 | 0.356 | -1.383 | 2018-03-29 | reject |
| 0015280201 | U | 28d (16 Mar) | 0.227 | 0.336 | -1.396 | | ✗ |
| 0015280201 | U | 56d (16 Feb) | 0.212 | 0.341 | -1.402 | 2018-04-05 | reject |
| 0015280201 | U | 49d (23 Feb) | 0.284 | 0.314 | -1.411 | 2018-04-12 | reject |
| 0015280201 | U | 56d (16 Feb) | 0.146 | 0.362 | -1.440 | 2018-04-05 | reject |
| 0015281401 | L | 17d (27 Mar) | 0.218 | 0.333 | -1.474 | 2018-04-12 | reject |
| 0015280201 | U | 84d (19 Jan) | 0.165 | 0.351 | -1.490 | 2018-03-01 | reject |
| 0015280201 | U | 77d (26 Jan) | 0.153 | 0.355 | -1.495 | 2018-03-15 | reject |
| 0015280201 | U | 84d (19 Jan) | 0.145 | 0.356 | -1.519 | 2018-03-01 | reject |
| 0017241701 | A | 65d (07 Feb) | 0.260 | 0.313 | -1.532 | 2018-03-29 | reject |
| 0015280201 | U | 7d (06 Apr) | 0.159 | 0.348 | -1.555 | | ✗ |
| 0015281401 | L | 39d (05 Mar) | 0.153 | 0.338 | -1.706 | 2018-04-05 | reject |
| 0015281401 | L | 39d (05 Mar) | 0.133 | 0.341 | -1.759 | 2018-04-05 | reject |
| 0015280201 | U | 23d (21 Mar) | 0.222 | 0.285 | -2.051 | | ✗ |
| 0015280201 | U | 28d (16 Mar) | 0.230 | 0.266 | -2.250 | | ✗ |
| 0015280201 | U | 7d (06 Apr) | 0.144 | 0.276 | -2.515 | | ✗ |
| 0015280201 | U | 44d (28 Feb) | 0.197 | 0.247 | -2.635 | 2018-04-12 | reject |
| 0015281401 | L | 39d (05 Mar) | 0.135 | 0.255 | -2.816 | 2018-04-05 | reject |

auto | preview | send



# What we expect?

| prediction model | N | only P | only S | P + S |
|---|---|---|---|---|
| uninformed selection | 60% | 13% | 21% | 6% |
| screening model (top 150 patients) | 29% | 30% | 18% | 23% |
| selection best (CHECK 2y shift, top 150) | 35% | 31% | 21% | 14% |
| selection worst (MUST 5y shift, top 60) | 55% | 32% | 7% | 7% |

*(optimistic estimate on CHECK data)

- Could we do better with more patients / better data?

# What we expect?

| prediction model | N | only P | only S | P + S |
|---|---|---|---|---|
| uninformed selection | 60% | 13% | 21% | 6% |
| screening model (top 150 patients) | 29% | 30% | 18% | 23% |
| selection best (CHECK 2y shift, top 150) | 35% | 31% | 21% | 14% |
| selection worst (MUST 5y shift, top 60) | 55% | 32% | 7% | 7% |

*(optimistic estimate on CHECK data)

- Could we do better with more patients / better data?

PART IV

# Look into the Future

# Two battles of the future

## Battle for data

- tragedy of electronic records (originally set up for billing purposes)
- single point of storage
- who owns the data?
- who protects the data?

## Battle for doctors time

- administrative tasks vs. patient interactions
- will doctors become data clerks?
- who will make the tough calls?

# Two battles of the future

## Battle for data

- tragedy of electronic records
  (originally set up for billing purposes)
- single point of storage
- who owns the data?
- who protects the data?

## Battle for doctors time

- administrative tasks vs. patient interactions
- will doctors become data clerks?
- who will make the tough calls?

# Two battles of the future

## Battle for data

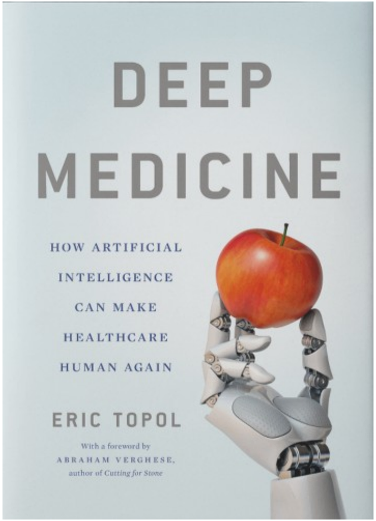
- tragedy of electronic records
  (originally set up for billing purposes)
- single point of storage
- who owns the data?
- who protects the data?

## Battle for doctors time

- administrative tasks vs. patient interactions
- will doctors become data clerks?
- who will make the tough calls?

Do airline pilots manage the ticket counter?
Do they count the passengers and run the safety demo?

# Worth fighting for?




## Key observation

- progressing erosion of doctor-patient relationship (distraction + burn out)

## Is AI a solution?

- AI can transform **everything** doctors do: note taking, medical scans, diagnosis, and treatment (cutting down the cost of medicine)
- by freeing physicians time, AI can create space for the real **human connection**, a doctor who **listens** and a patient who is **heard**

# References


Finlayson, S. G., Kohane, I. S., and Beam, A. L. (2019).
Adversarial attacks against medical deep learning systems.
*Computing Research Repository*, arXiv:1804.05296v3.

Gatys, L. A., Ecker, A. S., and Bethge, M. (2015).
A neural algorithm of artistic style.
*Computing Research Repository*, arXiv:1508.06576v2.

Wang, A., Pruksachatkun, Y., Nangia, N., Singh, A., Michael, J., Hill, F., Levy, O., and Bowman, S. R. (2019).
SuperGLUE: A stickier benchmark for general-purpose language understanding systems.
*Computing Research Repository*, arXiv:1905.00537v2.