

# Outils mathématiques et d'intelligence artificielle pour la prévention du cancer du sein

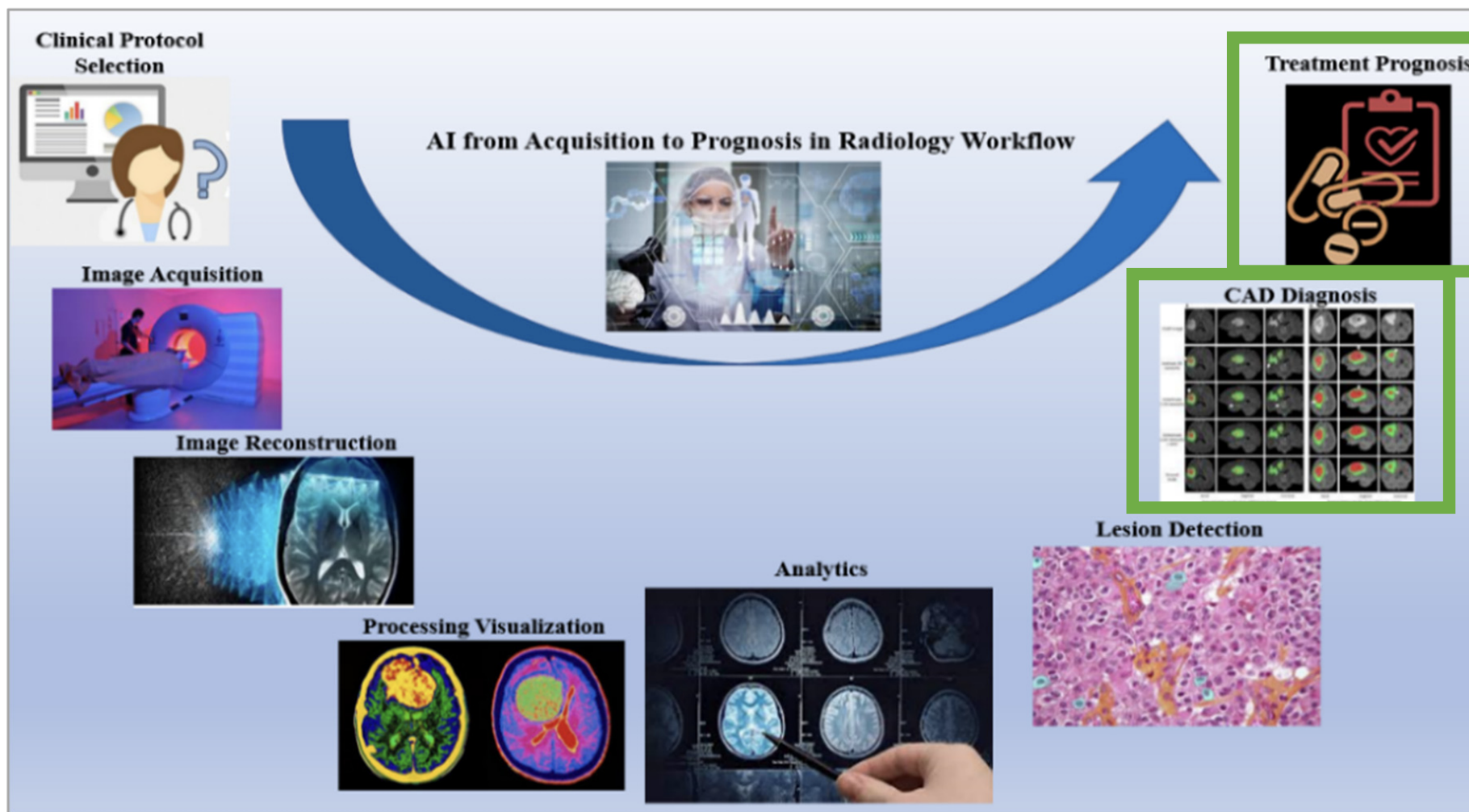
ARNAUD ATTYÉ, MD, PhD  
PIERRE-ETIENNE HEUDEL, MD



# DISCLOSURES

- GeodAlsics: DG
- GeodAlsics développe des solutions logicielles en médecine basées sur une technologie d'IA propriétaire, en médecine préventive et pronostique.

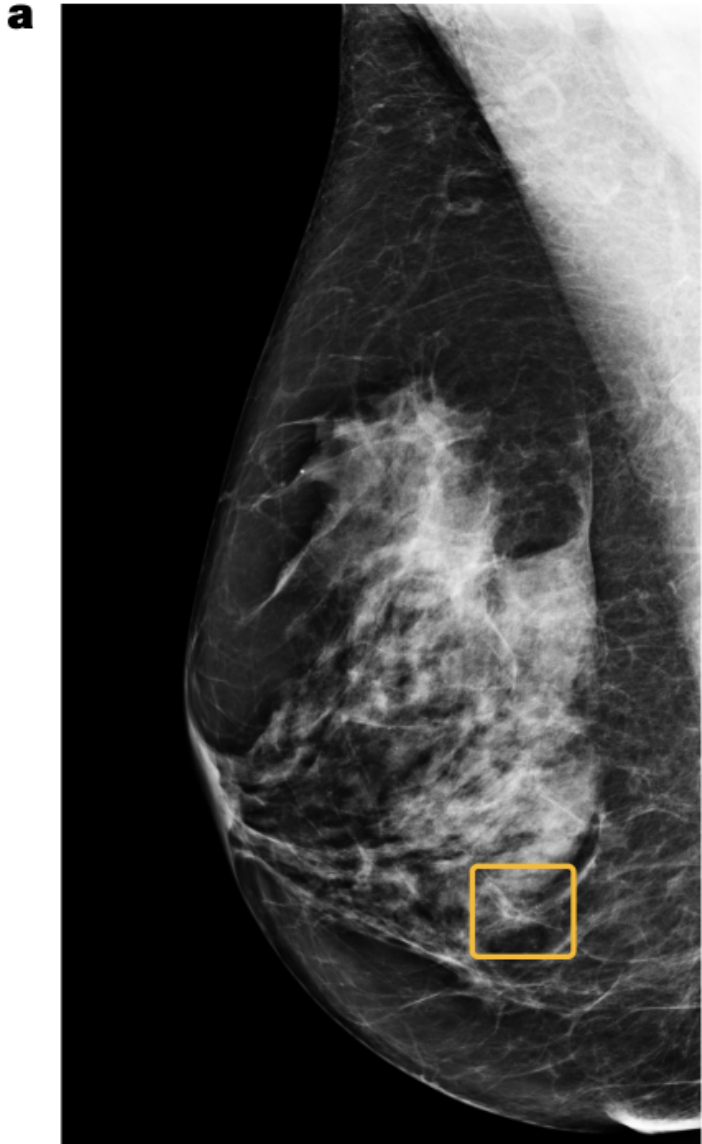
# DISCLOSURES



**GeodAlsics**

From Khanna et al. Healthcare 2022, 10, 2493.

# INTRO




nature

[Explore content](#) [About the journal](#) [Publish with us](#) [Subscribe](#)

[nature](#) > [articles](#) > [article](#)

Article | Published: 01 January 2020

## International evaluation of an AI system for breast cancer screening

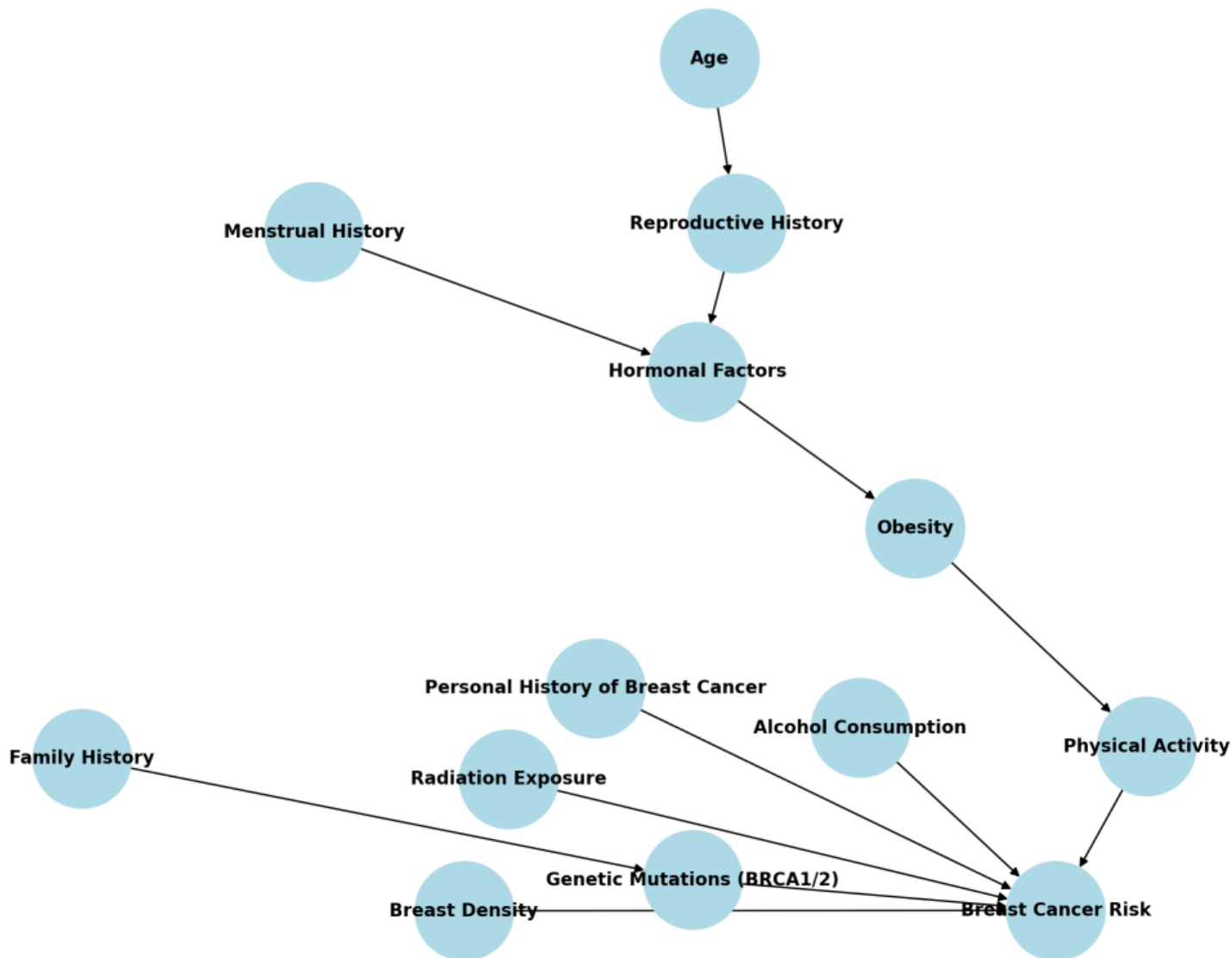
[Scott Mayer McKinney](#) , [Marcin Sieniek](#), [Varun Godbole](#), [Jonathan Godwin](#), [Natasha Antropova](#), [Hutan Ashrafian](#), [Trevor Back](#), [Mary Chesus](#), [Greg S. Corrado](#), [Ara Darzi](#), [Mozzlyar Etemadi](#), [Florencia Garcia-Vicente](#), [Fiona J. Gilbert](#), [Mark Halling-Brown](#), [Demis Hassabis](#), [Sunny Jansen](#), [Alan Karthikesalingam](#), [Christopher J. Kelly](#), [Dominic King](#), [Joseph R. Ledsam](#), [David Melnick](#), [Hormuz Mostofi](#), [Lily Peng](#), [Joshua Jay Reicher](#), ... [Shravya Shetty](#)  [+ Show authors](#)

[Nature](#) **577**, 89–94 (2020) | [Cite this article](#)

**106k** Accesses | **1512** Citations | **3895** Altmetric | [Metrics](#)

# OUTILS STATISTIQUES D'ÉVALUATION DE RISQUE

Acyclic Graph for Estimating Breast Cancer Risk Factors



*Outils basés sur de l'épidémiologie  
des tests génétiques  
de la biologie de soin courant  
des biomarqueurs d'imagerie*

*SELECTION DE VARIABLES A PRIORI SUR LA  
LITTERATURE*

*ChatGPT-o1*

# OUTILS STATISTIQUES D'ÉVALUATION DE RISQUE

## Score d'Eisinger



### Branche Maternelle

- 5 - Mutation BRCA1 ou 2 identifiée dans la famille.
- 4 - Cancer du sein chez une femme < 30 ans.
- 3 - Cancer du sein chez une femme entre 30 et 39 ans.
- 2 - Cancer du sein chez une femme 40 et 49 ans.
- 1 - Cancer du sein chez une femme 50 et 70 ans.
- 4 - Cancer du sein chez un homme.
- 3 - Cancer de l'ovaire.

### Branche Paternelle

- 5 - Mutation BRCA1 ou 2 identifiée dans la famille.
- 4 - Cancer du sein chez une femme < 30 ans.
- 3 - Cancer du sein chez une femme entre 30 et 39 ans.
- 2 - Cancer du sein chez une femme 40 et 49 ans.
- 1 - Cancer du sein chez une femme 50 et 70 ans.
- 4 - Cancer du sein chez un homme.
- 3 - Cancer de l'ovaire.

## IBIS Breast Cancer Risk Evaluation Tool

Description	Software Downloads	Documentation	Screenshots & Examples	Software Change Log
-------------	--------------------	---------------	------------------------	---------------------

### FAQs

**NEW! v8** [ZIP]



### Description of breast cancer risk program

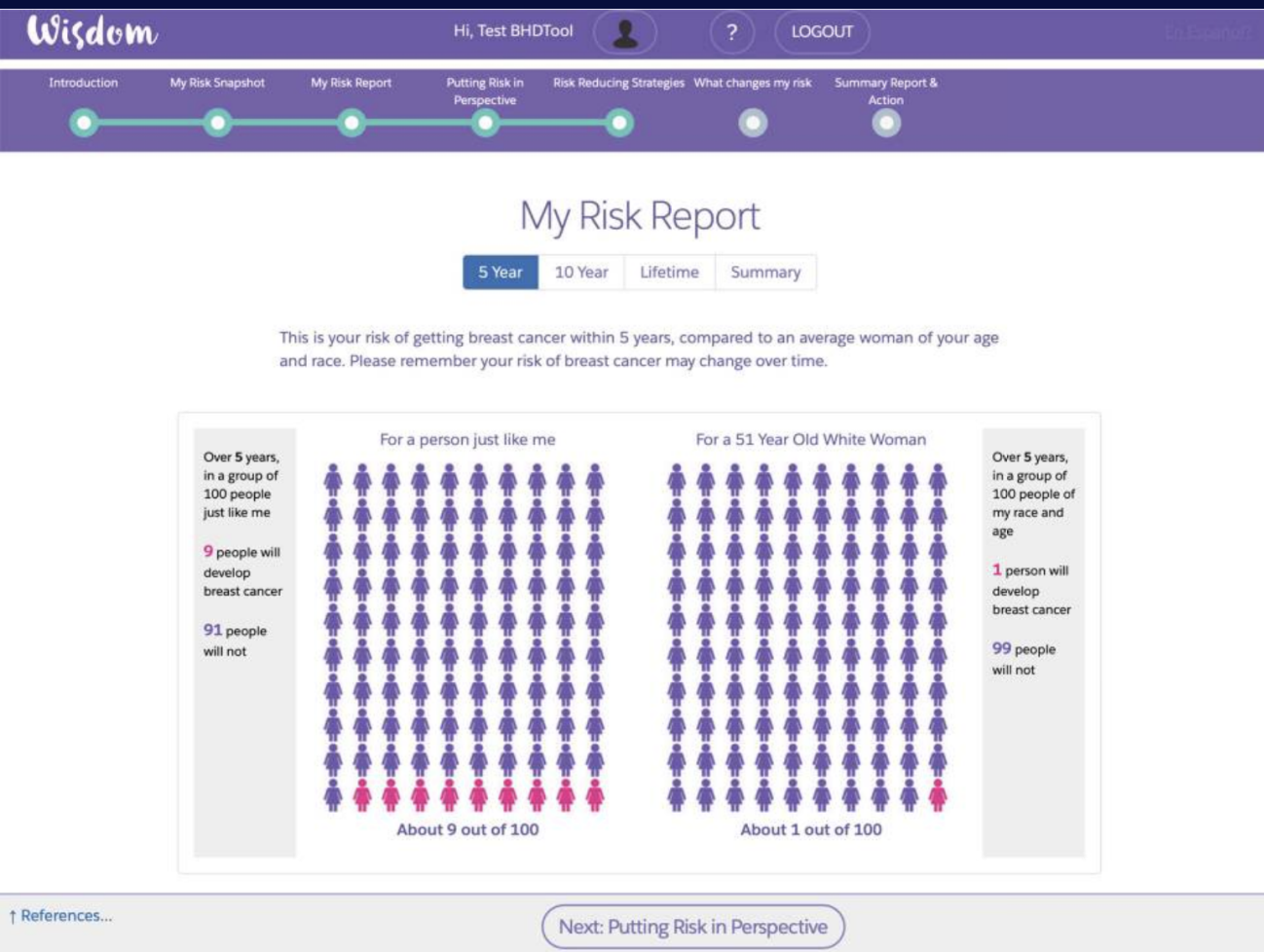
The program assumes that there is a gene predisposing to breast cancer in addition to the *BRCA1/2* genes. The woman's family history is used to calculate the likelihood of her carrying an adverse gene, which in turn affects her likelihood of developing breast cancer. The risks of developing breast cancer for the general population were taken from data on the first breast cancer diagnosis (ICD-10 code C50) in Thames Cancer Registry area (UK) between 2005-2009. The risk from family history (caused by the adverse genes) is modelled to fit the results in "Familial Breast and Ovarian Cancer: A Swedish Population-based Register Study, Anderson H et al., American Journal of Epidemiology 2000, 152: 1154-1163".

The risk from other classical factors including age at first child and benign disease are combined with familial risk.

The latest version of the model (v8) incorporates mammographic density.

## Modèle Tyrer-Cuzick

# OUTILS STATISTIQUES D'ÉVALUATION DE RISQUE



▶ NPJ Breast Cancer. 2021 Jun 17;7:78. doi: [10.1038/s41523-021-00288-8](https://doi.org/10.1038/s41523-021-00288-8)

## Development and pilot of an online, personalized risk assessment tool for a breast cancer precision medicine trial

[Holly Keane](#)<sup>1,2,#</sup>, [Yash S Huilgol](#)<sup>3,4,#</sup>, [Yiwey Shieh](#)<sup>3</sup>, [Jeffrey A Tice](#)<sup>3</sup>, [Jeff Belkora](#)<sup>5</sup>, [Karen Sepucha](#)<sup>6</sup>, [W Patrick Shibley](#)<sup>3</sup>, [Tianyi Wang](#)<sup>1</sup>, [Mandy Che](#)<sup>1</sup>, [Deborah Goodman](#)<sup>7</sup>, [Elissa Ozanne](#)<sup>8</sup>, [Allison Stover Fiscalini](#)<sup>1</sup>, [Laura J Esserman](#)<sup>1,✉</sup>

▶ Author information ▶ Article notes ▶ Copyright and License information

PMCID: PMC8211836 PMID: [34140528](https://pubmed.ncbi.nlm.nih.gov/34140528/)

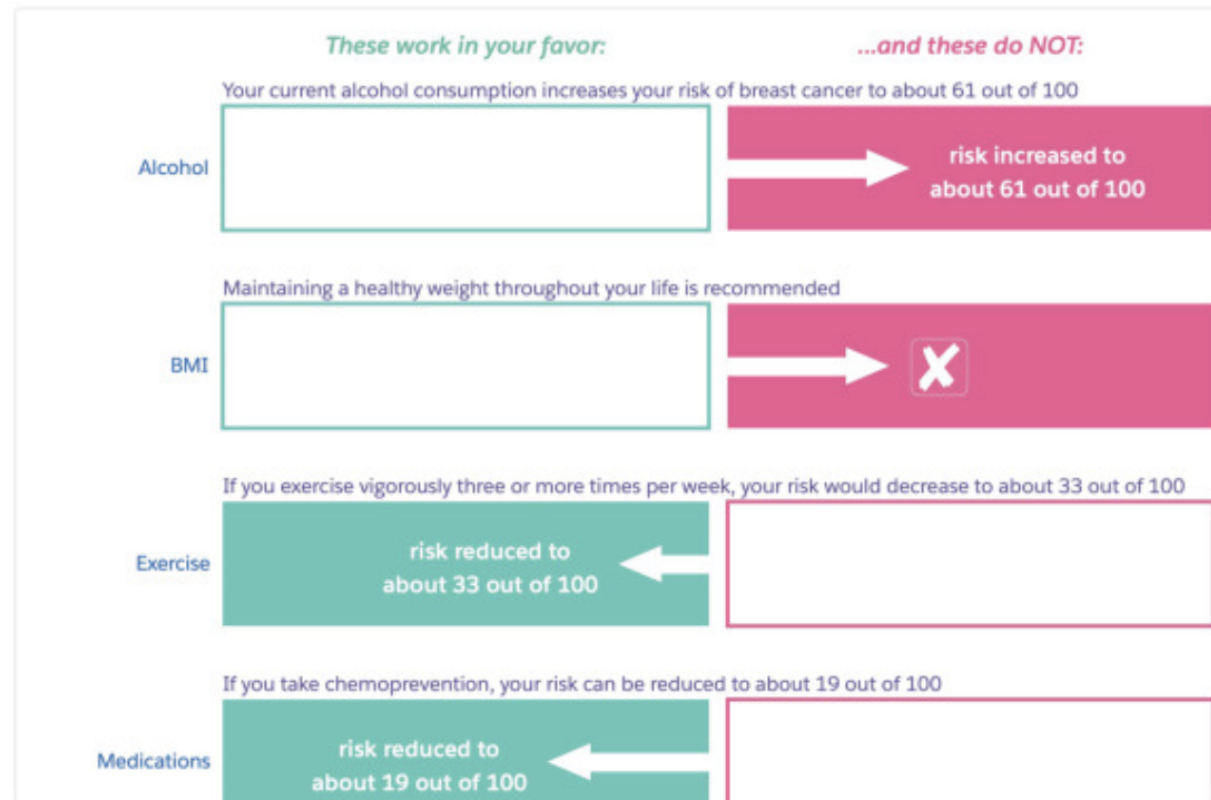
# OUTILS STATISTIQUES D'ÉVALUATION DE RISQUE



## Exploring what changes my risk

The risk score we provided includes factors such as your age, race, family history of breast cancer, breast biopsy history, and breast density. Your lifetime risk score is about **38** out of 100.

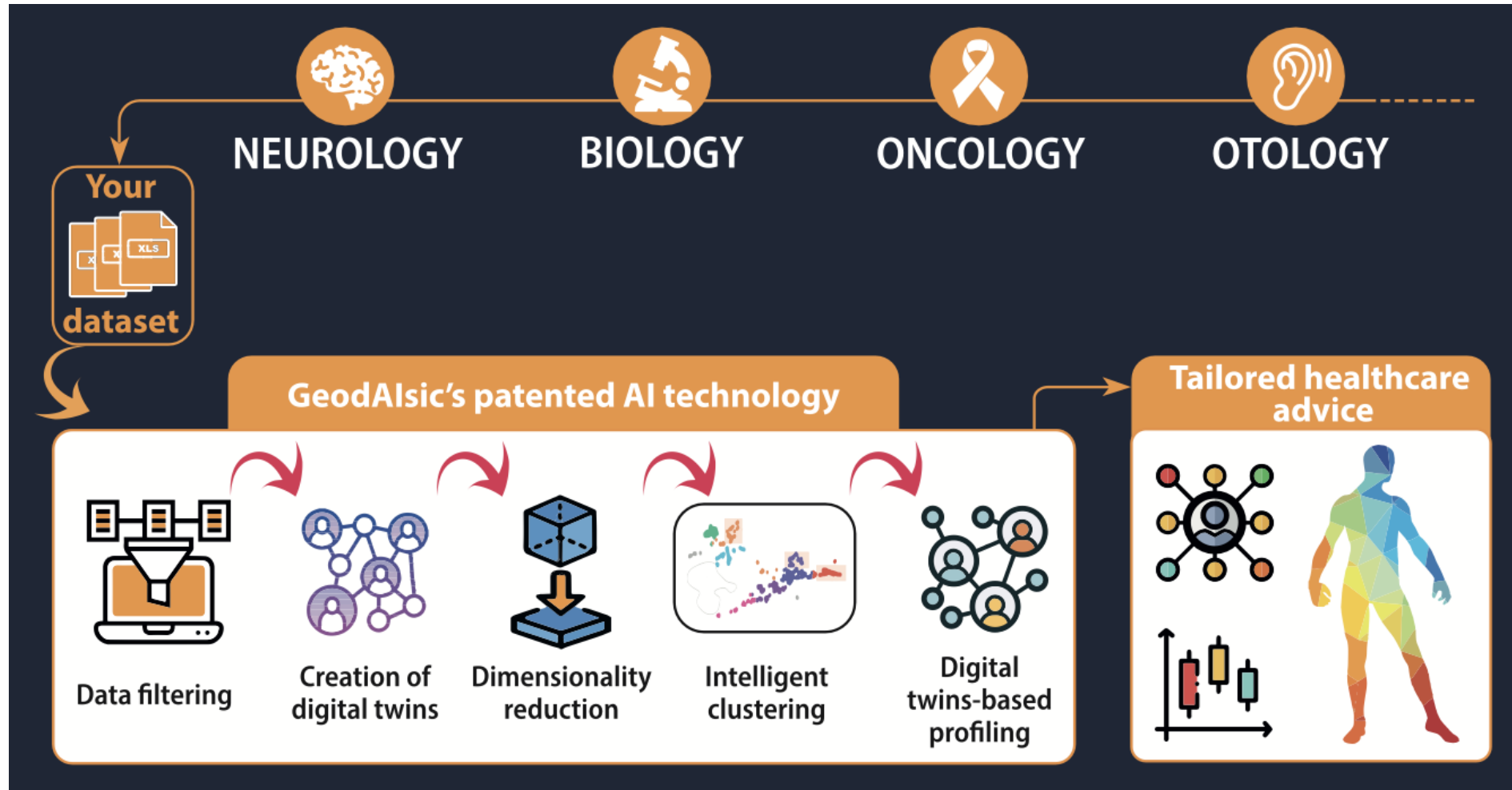
However, risk-reducing strategies can also increase or decrease your calculated personalized risk as seen below.



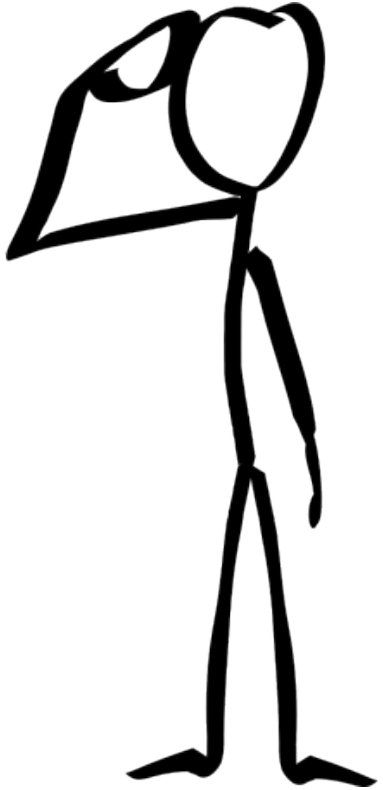
# OUTILS STATISTIQUES D'ÉVALUATION DE RISQUE

- **Avantage:** modèles statistiques bien maîtrisés.
- **Inconvénients:** limités par la malédiction de la dimension, pas de gestion des variables conjointes, difficulté de gérer les variables continues.

# APPRENTISSAGE MACHINE NON SUPERVISÉ



# APPRENTISSAGE MACHINE NON SUPERVISÉ



Facteurs  
biologiques:  
Vitamine D/  
oestrogène/  
insuline à jeun

Origine ethnique

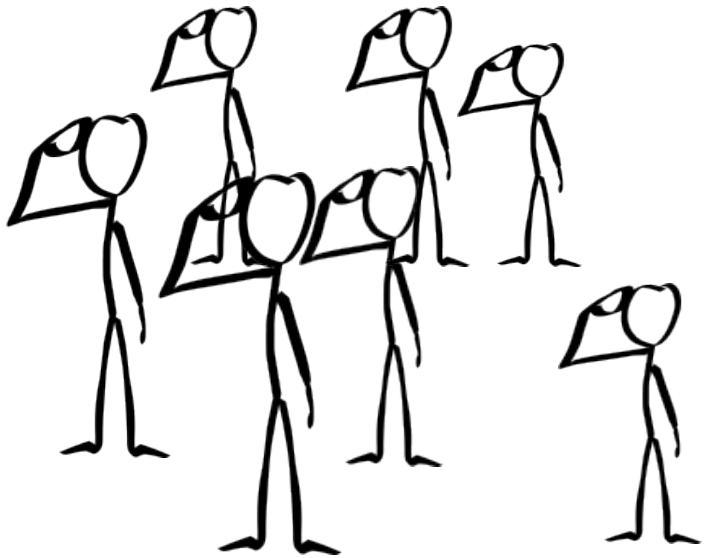
Age du premier  
enfant

Age civil

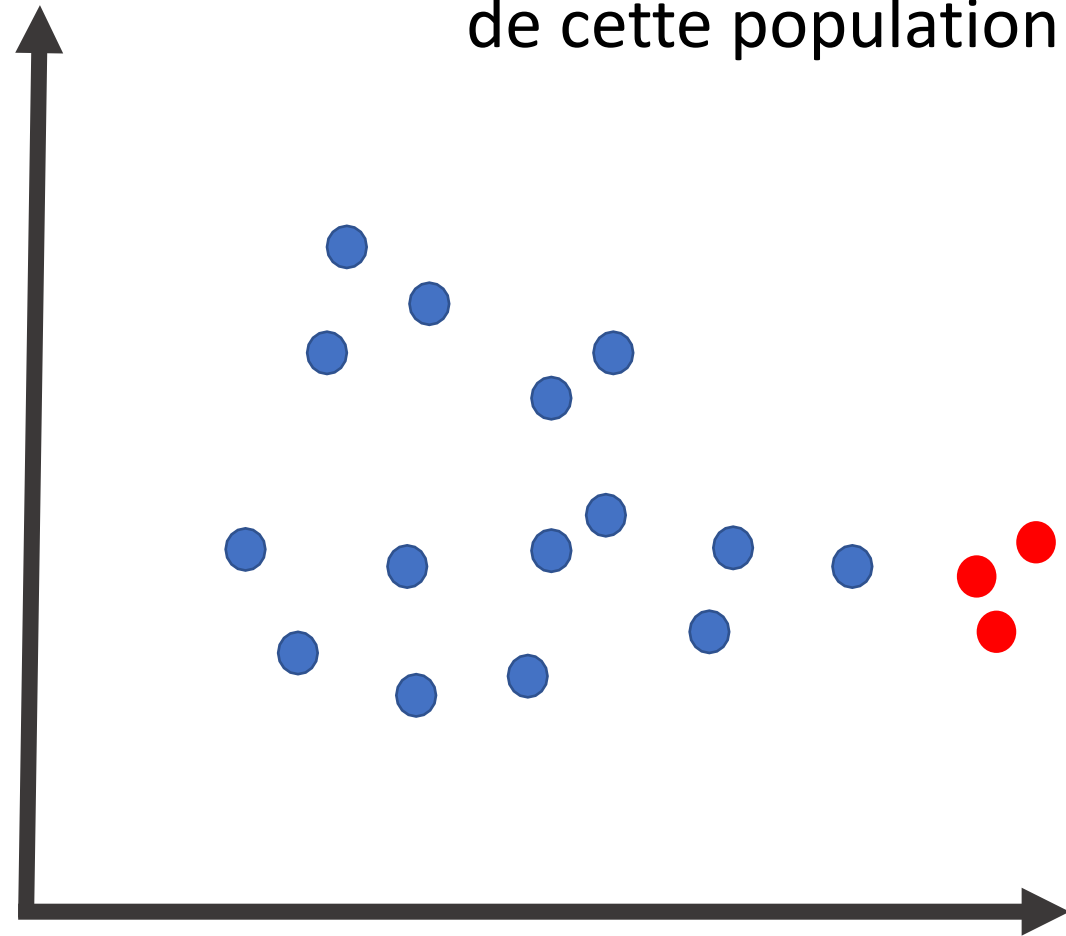
IA non  
supervisée

# APPRENTISSAGE MACHINE NON SUPERVISÉ

Population de  
sujets



Apprentissage machine  
de cette population



# APPRENTISSAGE MACHINE NON SUPERVISÉ

Nouveau patient

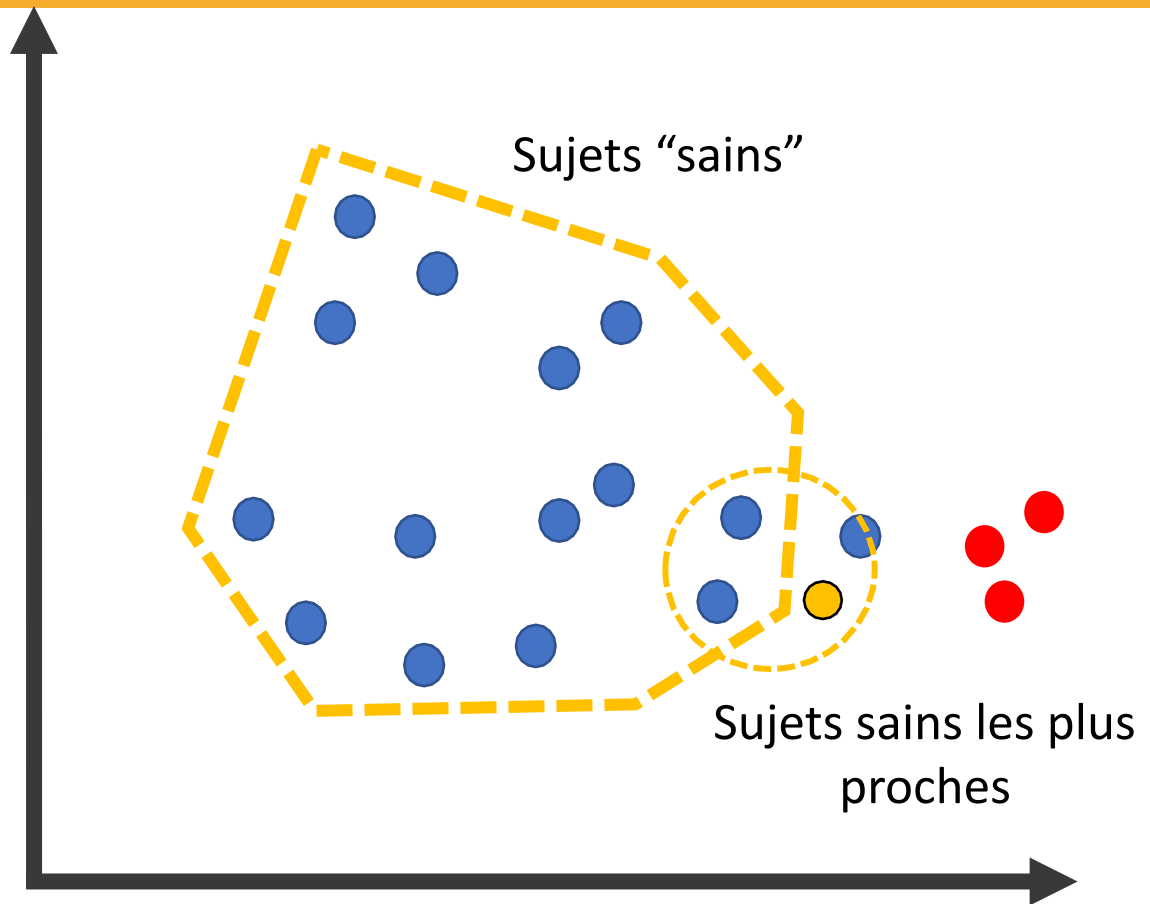


Projection dans la population

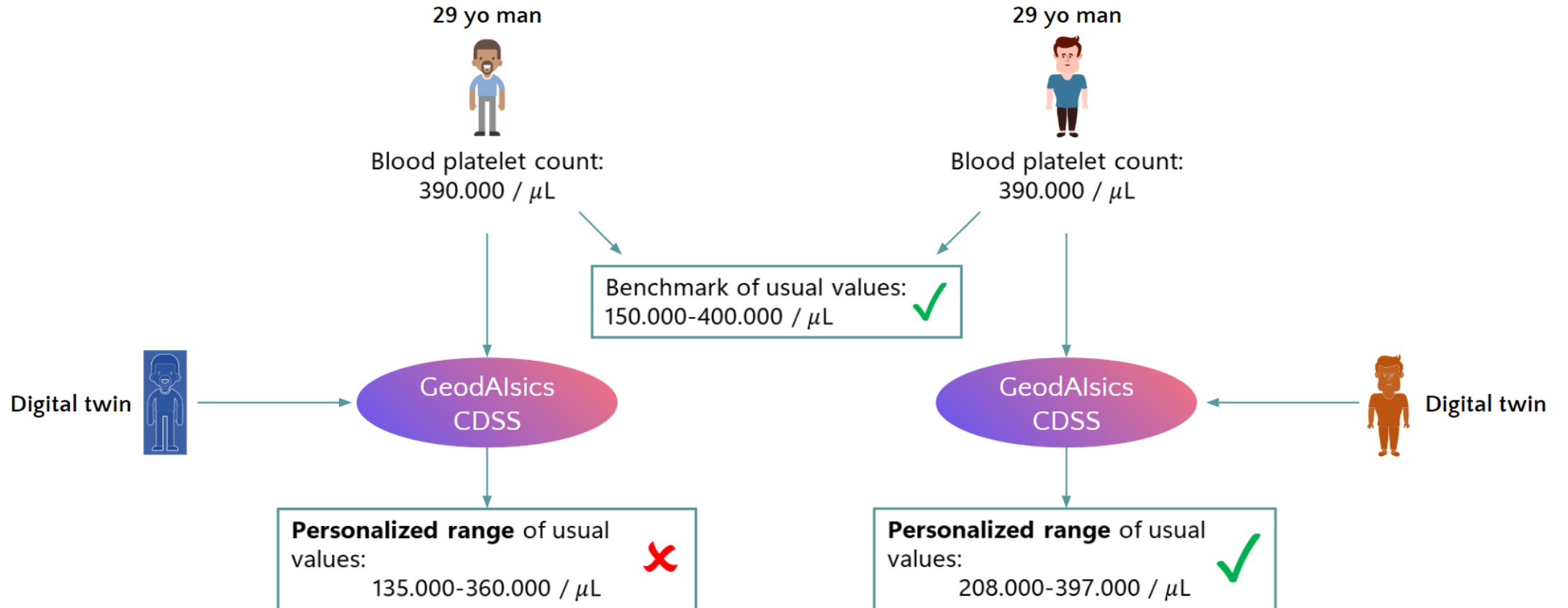


Création d'un jumeau numérique

C'est la comparaison entre le nouveau patient et son jumeau qui permet de détecter des signes précoces de risque de cancer du sein dans les données

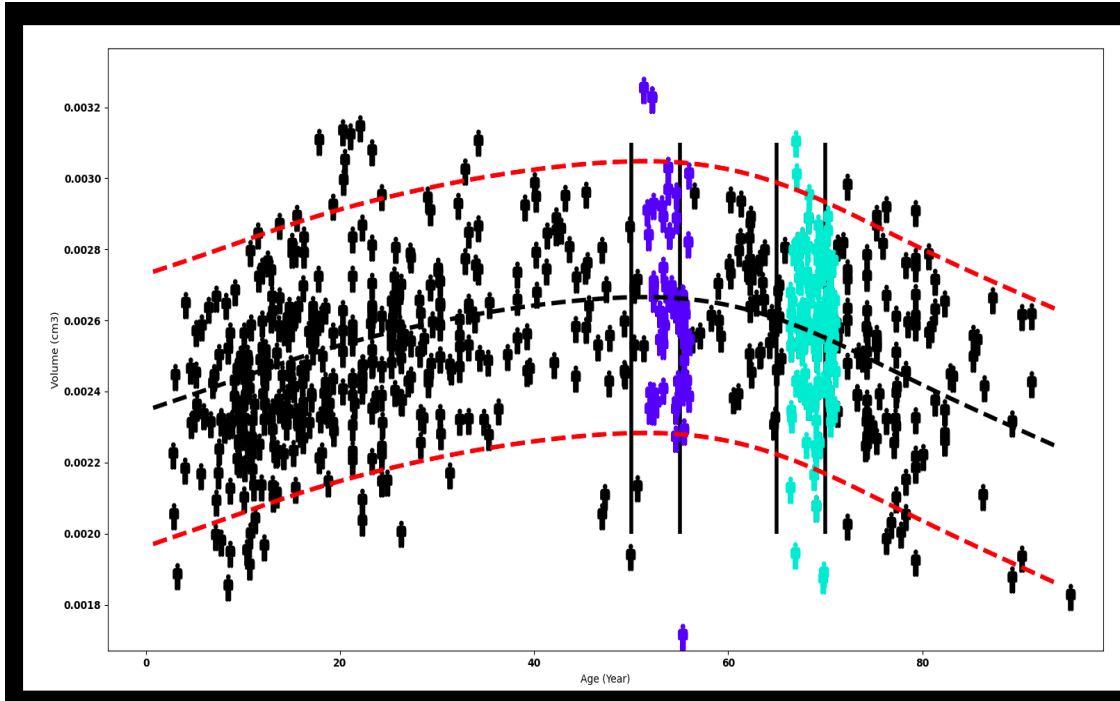


# APPRENTISSAGE MACHINE NON SUPERVISÉ

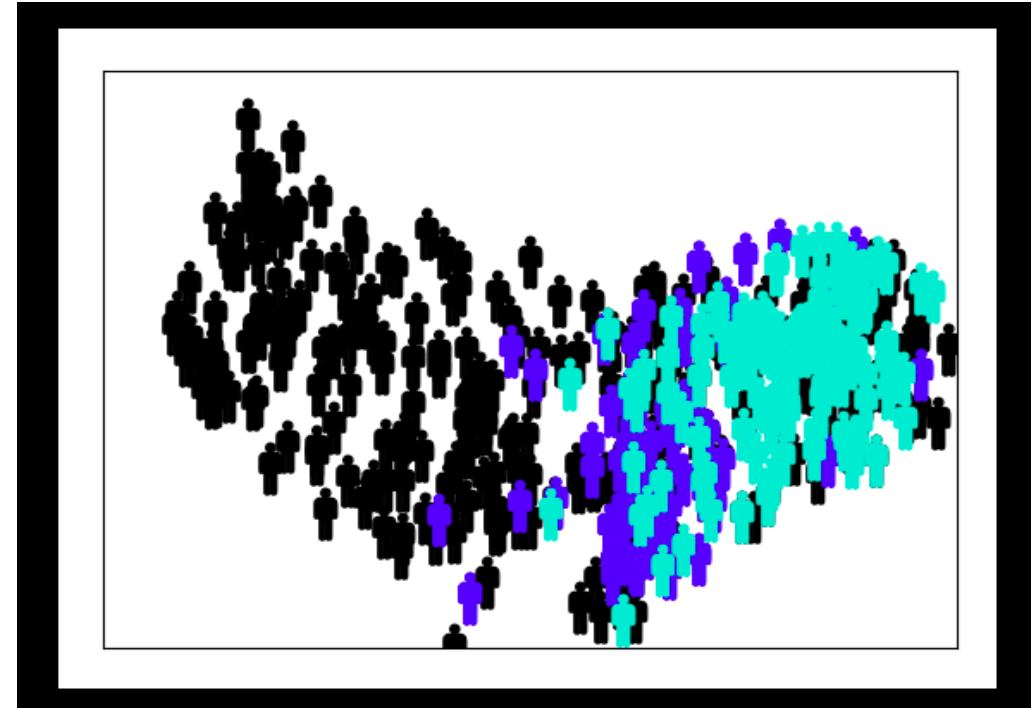


# APPRENTISSAGE MACHINE NON SUPERVISÉ

- Distribution et loi normale



Normes classiques univariés basées sur l'âge



Modélisation de la population basée sur l'ensemble des données

# APPRENTISSAGE MACHINE NON SUPERVISÉ

1

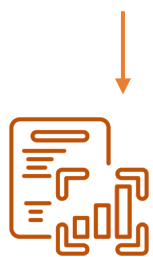
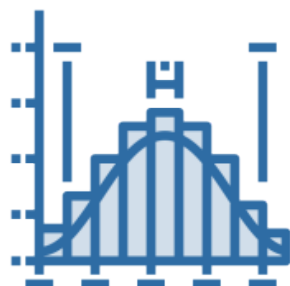
Etude de l'impact des variables sur le risque de cancer du sein



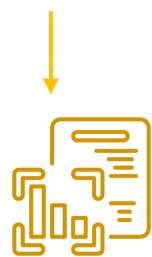
*SELECTION DE VARIABLES A PRIORI SUR LA LITTÉRATURE ET BASÉE SUR LES DONNÉES*

2

Plusieurs patients sont synthétisés avec des variables faibles et fortes



Profile 1



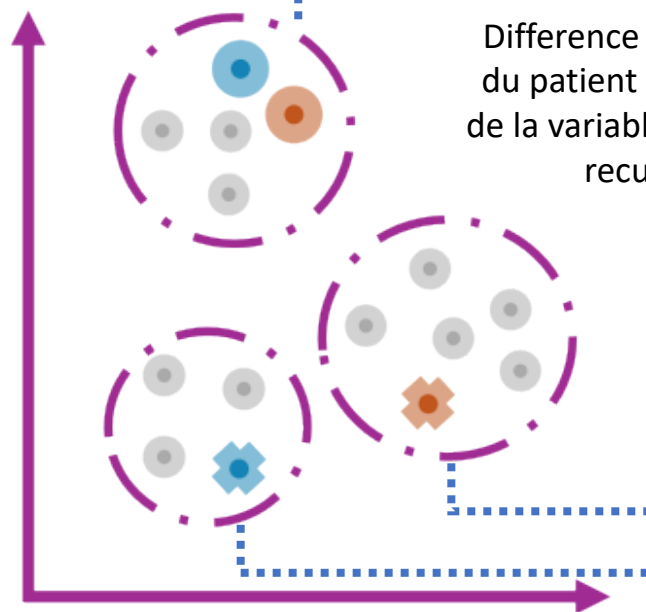
Profile 2

3

Etude de la position des patients dans les clusters

Pas de différence de position du patient en fonction de la variable: pas besoin dans le modèle

Différence de position du patient en fonction de la variable: variable à recueillir

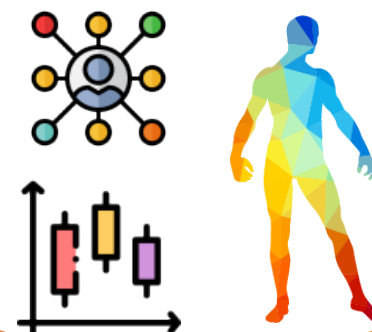


4

Sélection des variables indispensables

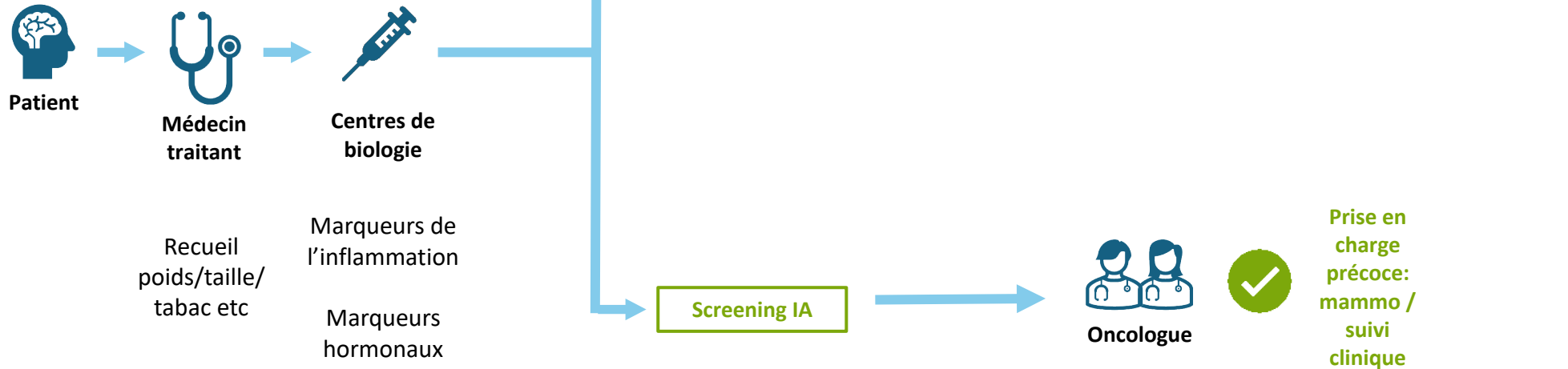


Prévention personnalisée pour le cancer du sein



# APPRENTISSAGE MACHINE NON SUPERVISÉ

- Sophie 37 ans,
- Pas d'antécédents familiaux significatifs
- Sophie n'est pas éligible au dépistage systématique par mammographie ni aux tests génétiques basés sur le score d'Eisinger.

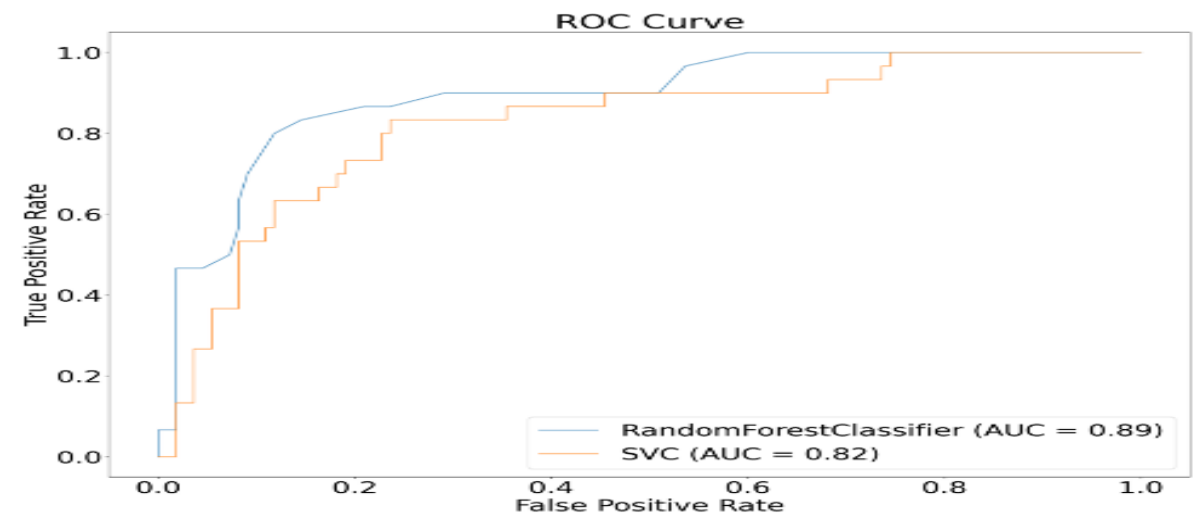
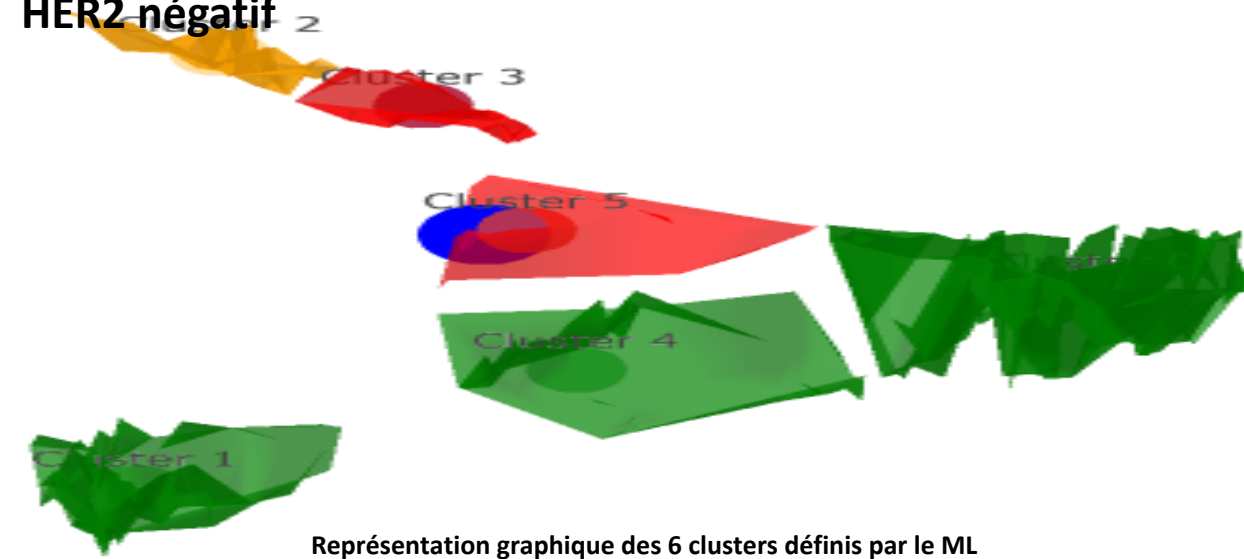


## PARCOURS DE SOIN SANS / AVEC IA

- Bilan de soin courant sans prise en charge?
- Bilan de fertilité?

# APPRENTISSAGE MACHINE NON SUPERVISÉ

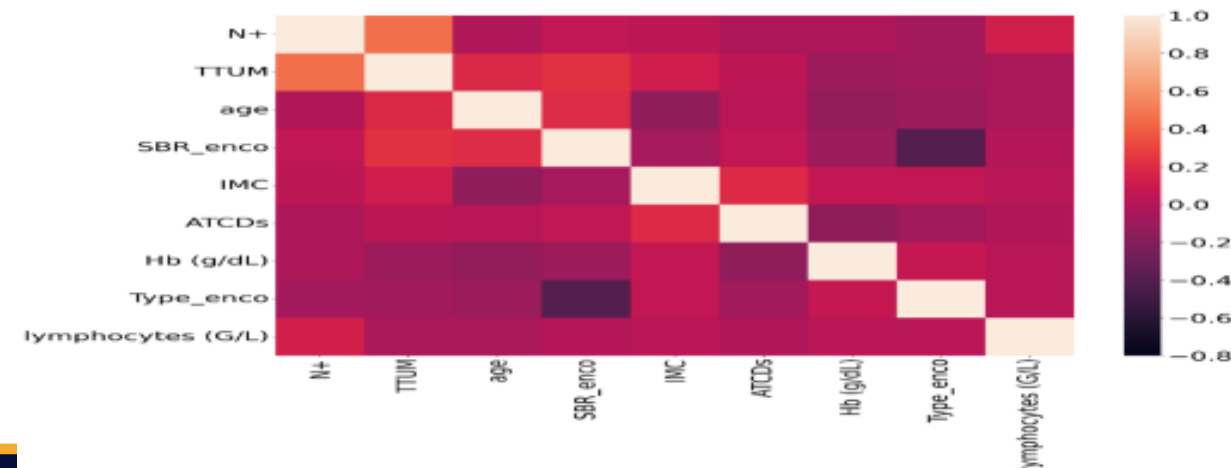
Optimisation de l'aide à la décision en oncogériatrie par l'utilisation des jumeaux numériques dans le cancer du sein localisé HER2 négatif



Courbes ROC pour les modèles prédictifs de mortalité à 5 ans.

Cluster	0	1	2	3	4	5
Hb (G/dL)	13.4	13.2	13.1	13	13.5	11.7
IMC	25	28.2	25.3	28.3	26	23.1
Nombre de N+	0.6	0.7	1	7.5	1	1.4
Tumeur (mm)	19.6	19.1	24.2	69.7	23.8	36.5
Age (années)	75.5	76.5	77.8	76.8	79.7	80.4
Lymphocytes (G/L)	1.8	2.1	1.9	1.6	3.6	1.5
Comorbidités	0	1	0.4	0.3	0	0.7
RH+/RH-	1	1	0	0.8	1	1
SBR High/low	0	0.1	0.7	0.5	0.7	0.7

Caractéristiques des 6 clusters définis par le ML



# APPRENTISSAGE MACHINE NON SUPERVISÉ

Hb (g/dL)

12.0 - +

IMC

25 - +

N+

2 - +

TTUM

30.0 - +

age

78 - +

lymphocytes (G/L)

1.0 - +

Presence of ATCDs

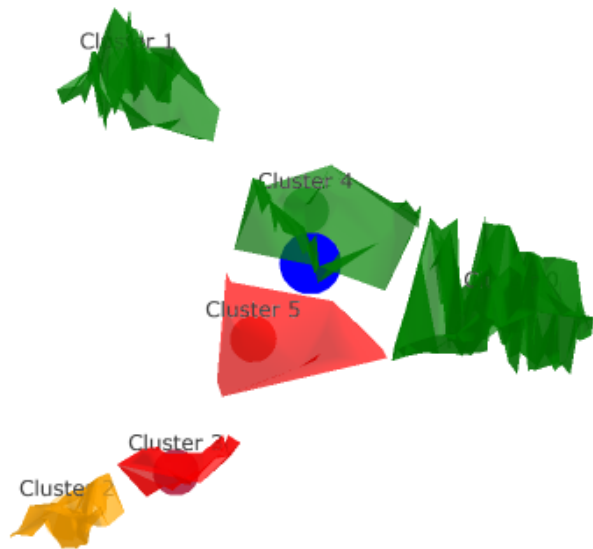
No ▾

Cancer Type

RH+ ▾

SBR

High (3) ▾



Hb (g/dL)

12.0 - +

IMC

25 - +

N+

2 - +

TTUM

30.0 - +

age

78 - +

lymphocytes (G/L)

0.5 - +

Presence of ATCDs

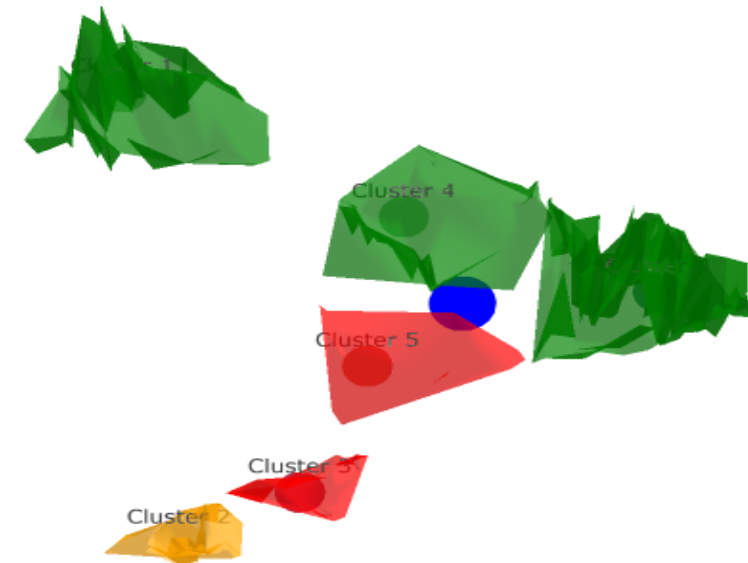
No ▾

Cancer Type

RH+ ▾

SBR

High (3) ▾

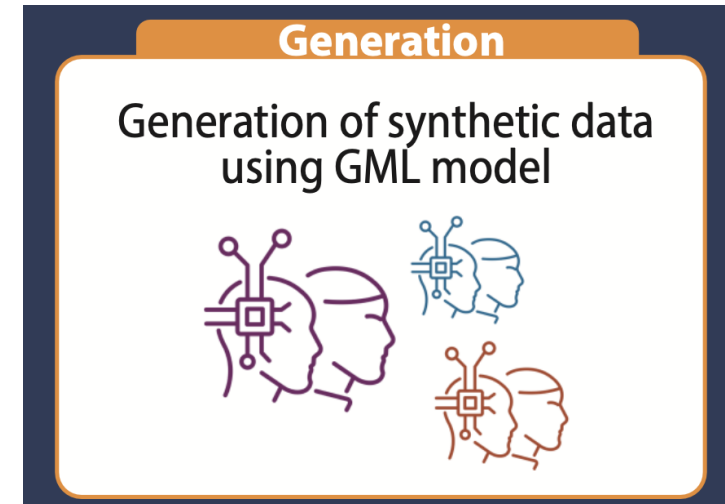
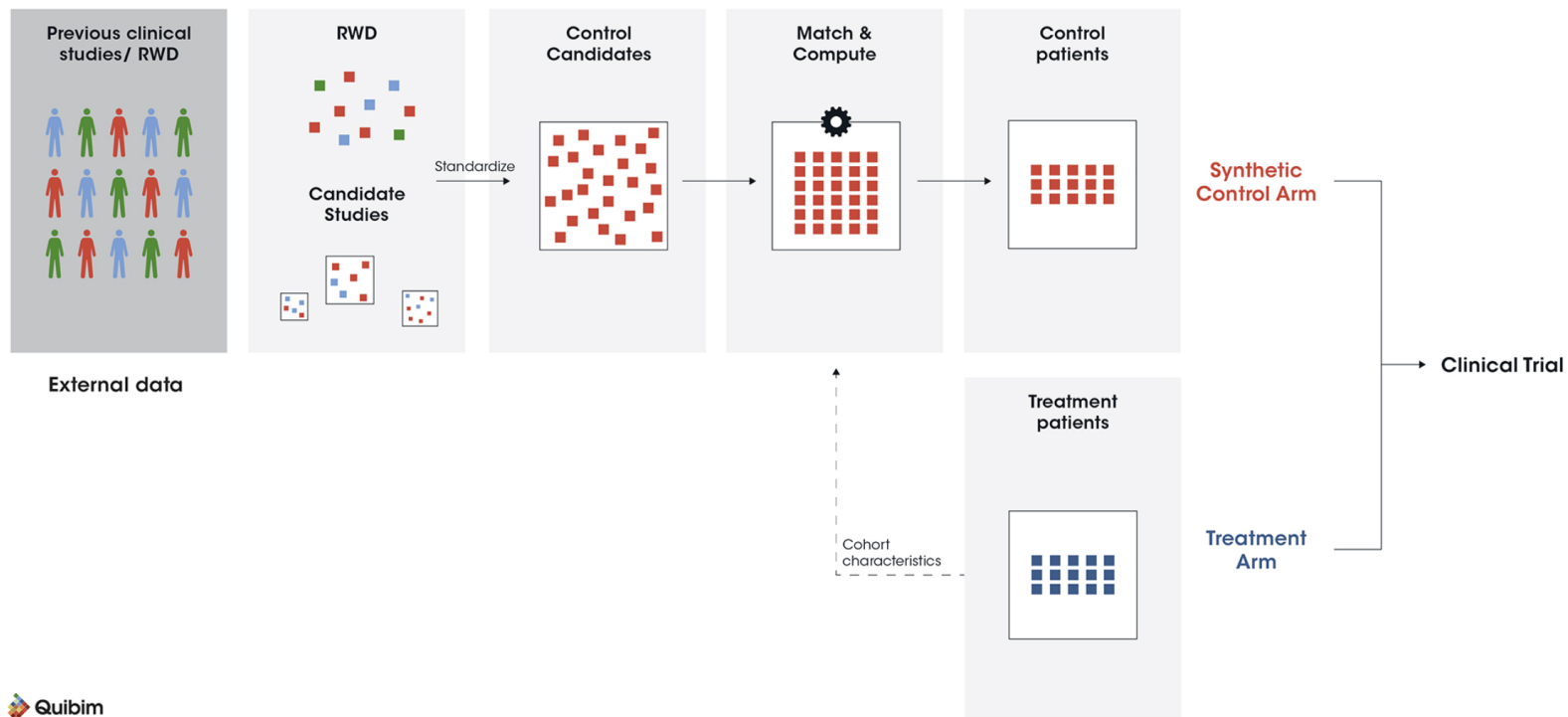


# SYNTHETIQUE

## Real World Data Synthetic Control Arm

## In silico medicine

How to create a synthetic control arm





EMA/FDA approved

To be validated

Article

## **Ethical Implications of Alzheimer's Disease Prediction in Asymptomatic Individuals through Artificial Intelligence**

Frank Ursin <sup>\*</sup>, Cristian Timmermann  and Florian Steger

Institute of the History, Philosophy and Ethics of Medicine, Ulm University, Parkstrasse 11, D-89073 Ulm, Germany; cristian.timmermann@uni-ulm.de (C.T.); florian.steger@uni-ulm.de (F.S.)

\* Correspondence: frank.ursin@uni-ulm.de; Tel.: +49-731-500-399-07

Les raisons de proposer des tests prédictifs à des individus asymptomatiques sont le droit de savoir, un bilan positif de l'évaluation bénéfico-risque, et l'opportunité de planifier pour l'avenir.

Les raisons contre sont le (manque de traitements modifiant la maladie), la précision et l'explicabilité des prédictions assistées par l'IA, le droit de ne pas savoir, et les menaces aux droits sociaux.

Les tests pré-symptomatiques ne devraient être proposés que sur demande afin d'éviter de causer du tort.

Nous recommandons de développer une formation pour les médecins dans la communication des prédictions assistées par l'IA.

# CONCLUSION

- Les nouveaux outils d'IA non supervisés, notamment en apprentissage de population, offre la possibilité d'une médecine de prévention personnalisée mais il est nécessaire de poursuivre l'évaluation des outils et de leurs impacts.
- Les biomarqueurs innovants (densité mammaire issues du deep learning, génétique, cellules tumorales circulantes...) ont une plus value...et un coût pour la société.
- L'idée des autorités régulatrices est de faire plus de prévention avec moins de dépenses. Et de mieux exploiter les biomarqueurs déjà disponibles.

# CONTACTS

[arnaud@geodaisics.com](mailto:arnaud@geodaisics.com)

