PR[AI]RIE

PaRis Artificial Intelligence Research InstitutE

Décision médicale : concepts et réalités

Anita Burgun

12 Mai 2021

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS

Université de Paris

Type	Catégory	Sub category
Gathering additional information	Decision to obtain information from other source	Ordering test, consulting colleague, seeking external information
Evaluating test result	Simple, normative assessments of clinical findings and tests	Positive, negative, ambiguous
Defining problem	Complex, interpretative assessments that define what the problem is and reflect a medically informed conclusion	Diagnostic conclusion, evaluation of health state, aetiological inference, prognostic judgement
Drug related	Decision to start, refrain from, stop, alter or maintain a drug regimen	Start, stop, alter, maintain, refrain
Therapeutic procedure related	Decision to intervene on a medical problem, plan, perform or refrain from therapeutic procedures of a medical nature	Start, stop, alter, maintain, refrain
Legal and insurance related	Medical decision concerning the patient, which is based on or restricted by legal regulations or financial arrangements	Sick leave, drug refund, insurance, disability
Contact related	Decision regarding admittance or discharge from hospital, scheduling of control and referral to other parts of the healthcare system	Admit, discharge, follow-up, referral
Advice and precaution	Decision to give the patient advice or precaution, thereby transferring responsibility for action from the provider to the patient	Advice, precaution

Decision strategies in medicine

The simplest case : interpretation of a test result

- A clinical examination is the execution of idealised tests normatively assessing bodily functions.
- Sensitivity / specificity of the test, ideally by comparison with a "**gold standard**."
- **Lack of such clear-cut gold standard**
  - Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. BMJ. 2020 May 12;369:m1808. doi: 10.1136/bmj.m1808. PMID: 32398230.
  - A systematic review of the accuracy of covid-19 tests reported false negative rates of between 2% and 29% (sensitivity 71-98%), based on negative RT-PCR tests which were positive on repeat testing
- How to interpret the result in a specific context.
- Decision Identification and Classification Taxonomy for Use in Medicine (DICTUM)
  - normative assessments of diagnostic tests are defined as decisions

Decision based on clinical expertise

- Individual decisions by individual clinicians
- Clinical expertise : the proficiency and judgement that individual clinicians acquire through clinical practice
- Clinical expertise includes the general basic skills of clinical practice as well as the experience of the individual practitioner.

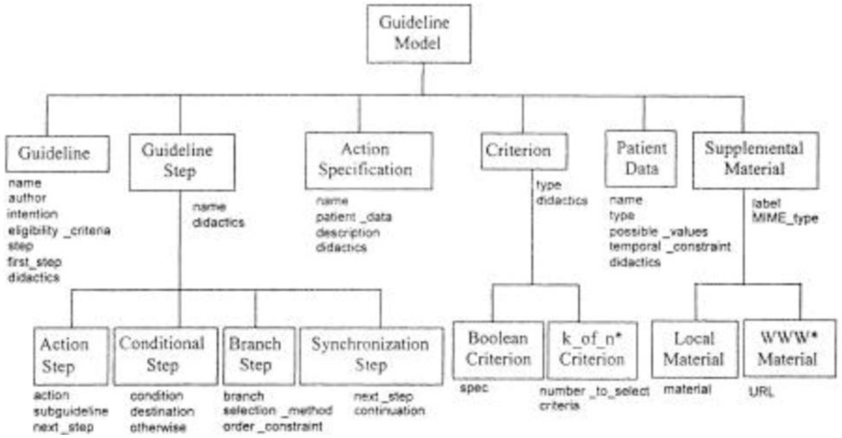
- Clinical expertise in the era of evidence based medicine and patient choice

- https://ebm.bmj.com/content/ebmed/7/2/36.full.pdf?fbclid=IwAR3TthF7vS1GuZdz6lld2cSIeTsk2cHsTLwjN0LTnsAWeVTI-VzlW1-Hzgg

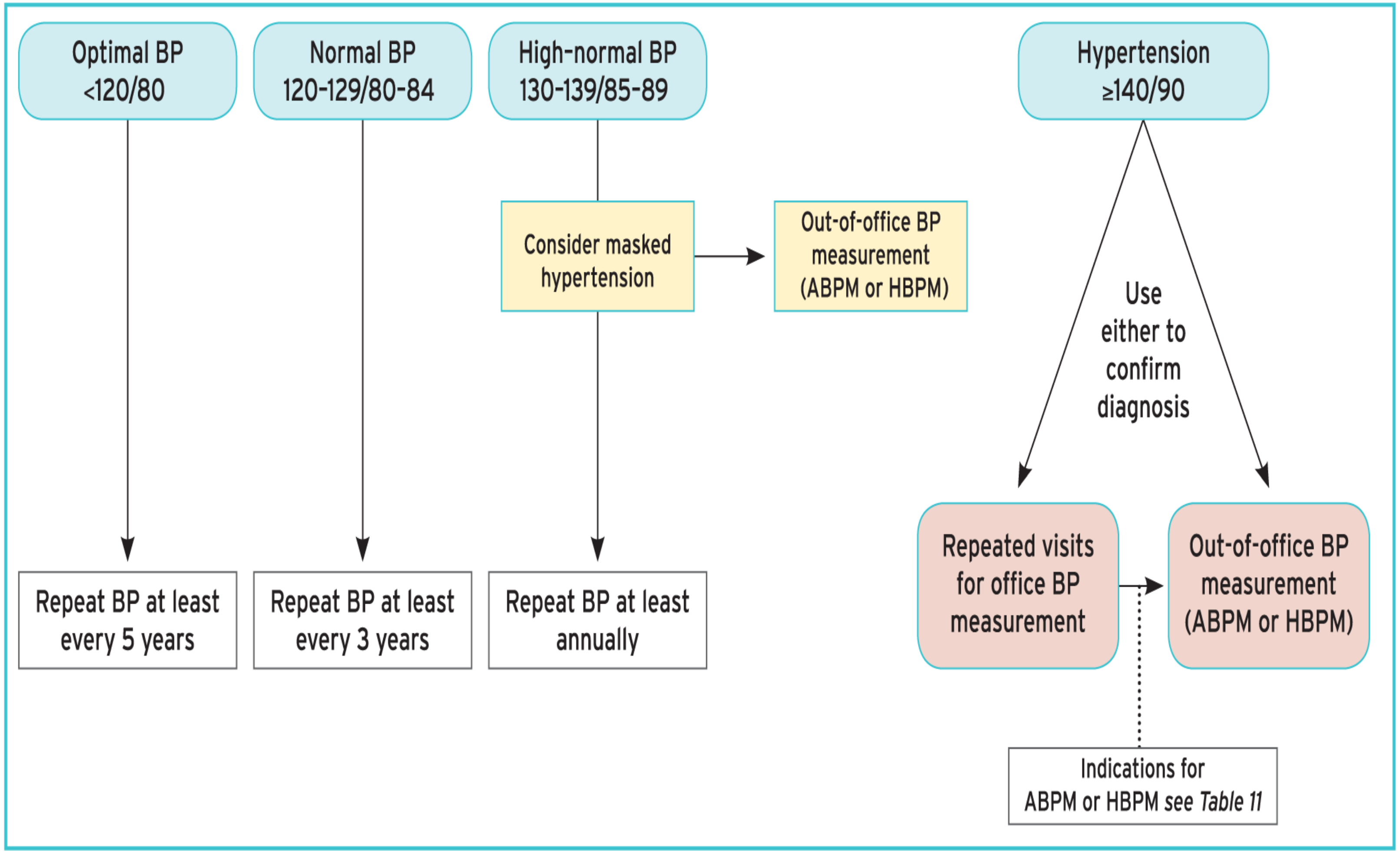
- Even excellent external evidence may be inapplicable to or inappropriate for an individual patient

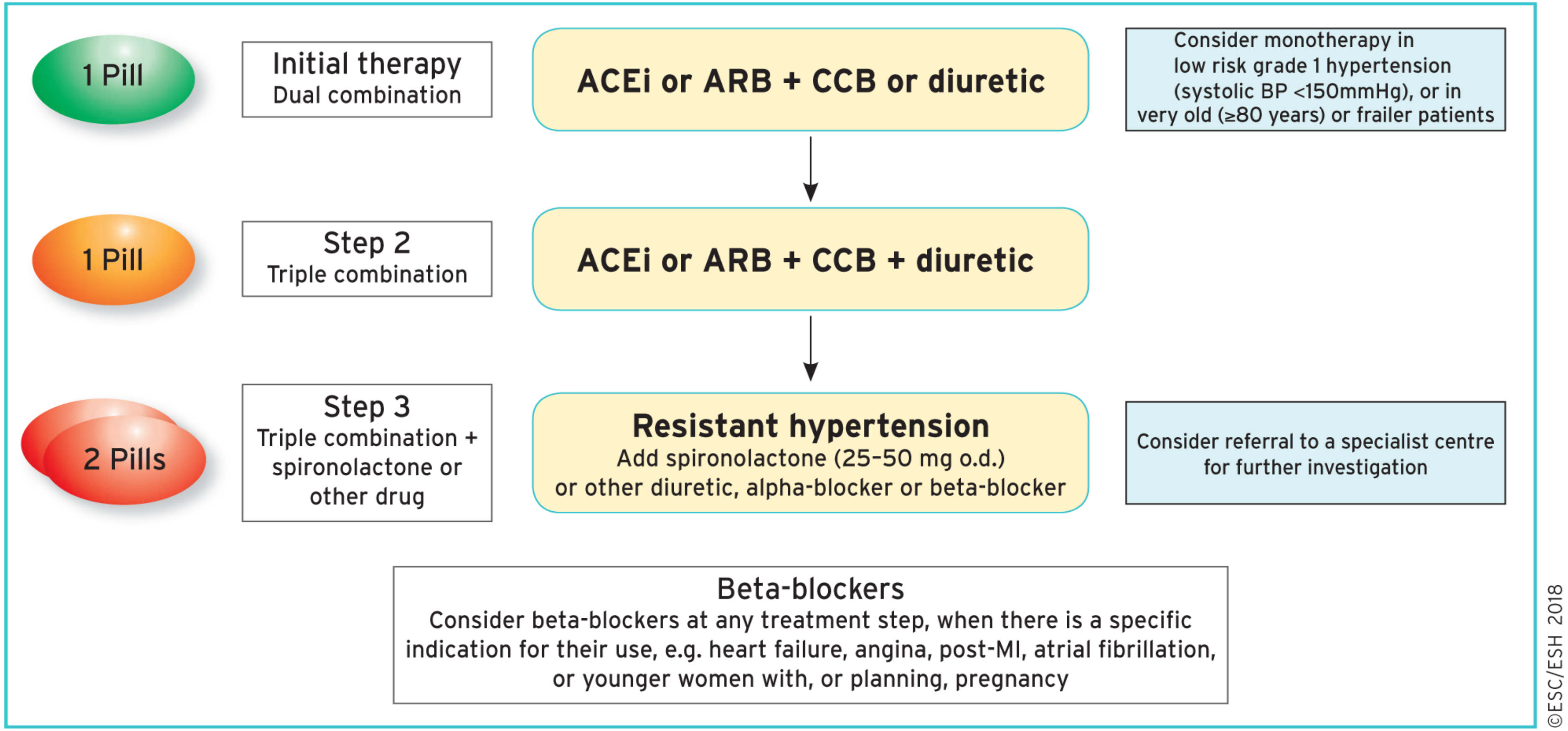
Standardized evidence-based decision

- Evidence Based Medicine (Sackett et al. BMJ, 1996)
- conscientious, explicit, and judicious use of current best (external) evidence (from well designed research) in making decisions about the care of individual patients
- Revue de littérature
- Consensus d'experts
- **Guidelines**



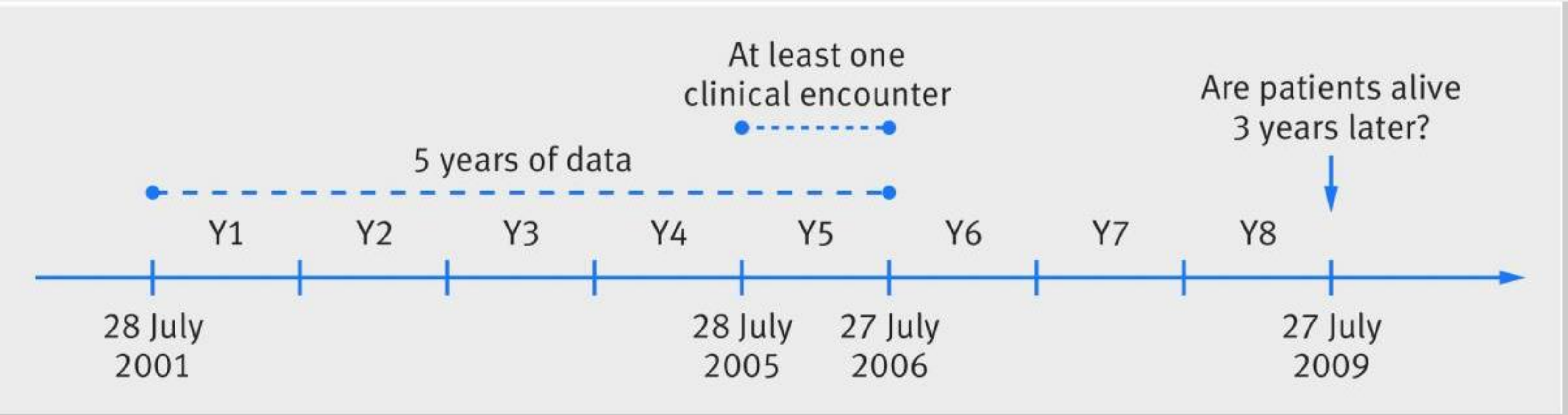
Ohno-Machado L, Gennari JH, Murphy SN, et al. The guideline interchange format: a model for representing guidelines. J Am Med Inform Assoc. 1998;5(4):357-372. doi:10.1136/jamia.1998.0050357




1 Pill
Initial therapy
Dual combination
ACEi or ARB + CCB or diuretic
Consider monotherapy in low risk grade 1 hypertension (systolic BP <150mmHg), or in very old (≥80 years) or frailer patients
1 Pill
Step 2
Triple combination
ACEi or ARB + CCB + diuretic
2 Pills
Step 3
Triple combination + spironolactone or other drug
Resistant hypertension
Add spironolactone (25-50 mg o.d.) or other diuretic, alpha-blocker or beta-blocker
Consider referral to a specialist centre for further investigation
Beta-blockers
Consider beta-blockers at any treatment step, when there is a specific indication for their use, e.g. heart failure, angina, post-MI, atrial fibrillation, or younger women with, or planning, pregnancy
©ESC/ESH 2018

Can decision be automated through machine learning?

- **DATA:**
- **2 hospitals**
- **Boston**
- 669 452 patients



- **MAIN OUTCOME MEASURES: Predictive value of 272 tests**
- **Time, day, value,**
- **RESULTS: fréquency predictive for 233 tests**

- Metadata were more predictive than results for118 tests

- processes & data in EHRs

Agniel D, Kohane IS, Weber GM. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. BMJ.2018 Apr 30;361:k1479.

Automated model versus treating physician
for predicting survival time of patients with metastatic cancer

Gensheimer M. et al.,, Journal of the American Medical Informatics Association, , ocaa290, https://doi.org/10.1093/jamia/ocaa290

- ML model to predict overall survival time using EHR (Epic, Verona, WI) data for patients seen for metastatic cancer in the Stanford Health Care system from 2008–2020.
- laboratory values, vital signs, ICD codes, CPT codes, text of provider notes and radiology reports, and medication administrations and prescriptions.
- Text e.g., oligometastatic state
- Compare with predictions made by the patient's radiation oncologist
- The ML model's survival predictions were more accurate than the physician's prediction.
- combining the ML model and physician's prediction resulted in a statistically significant improvement over the physician's prediction

- Open question : How shall we implement and use it?

World Health Organization

RESPONSIBLE
INNOVATION

Feature (+ means higher value increases survival)	
−Pulse	−Secondary malignant neoplasm of brain and spinal cord (ICD-9 198.3)
−Age	−Radiation treatment management (CPT 77427)
+Ephedrine (medication)	−Stereotactic MRI
+Complex radiation treatment delivery (CPT 77412)	−DNR/DNI order
+Office consultation (CPT 99244)	**−Encounter for palliative care (ICD-9 V66.7)**
+Out of bed to chair (nursing order)	**−Consult to palliative care**
+Red blood cell count	−Neoplasm-related pain (ICD-9 338.3)
+FDG PET/CT (skull to thighs)	
−Red cell distribution width (lab)	−MRI full spine with and without contrast
+Weight	

Batumalai V, et al. Pattern of palliative radiotherapy fractionation for brain metastases patients in New South Wales, Australia.
Radiother Oncol. 2020 Dec 22:S0167-8140(20)31246-9. doi: 10.1016/j.radonc.2020.12.020. Epub ahead of print. PMID: 33359268.

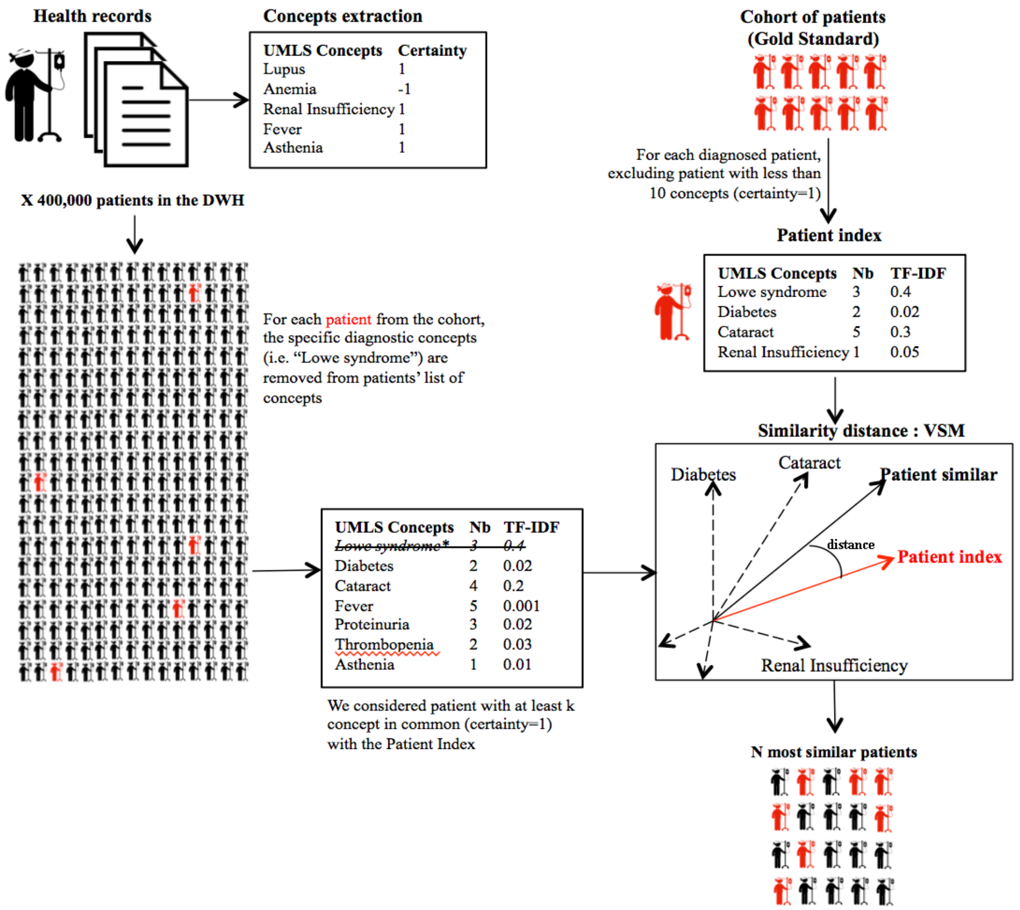
Can decision be automated through machine learning?

Frankovich J, Longhurst CA, Sutherland SM.
Evidence-based medicine in the EMR era.
N Engl J Med. 2011 Nov 10;365(19):1758-9.

Use of an EHR repository to inform a decision about anticoagulation in a patient with SLE.

11

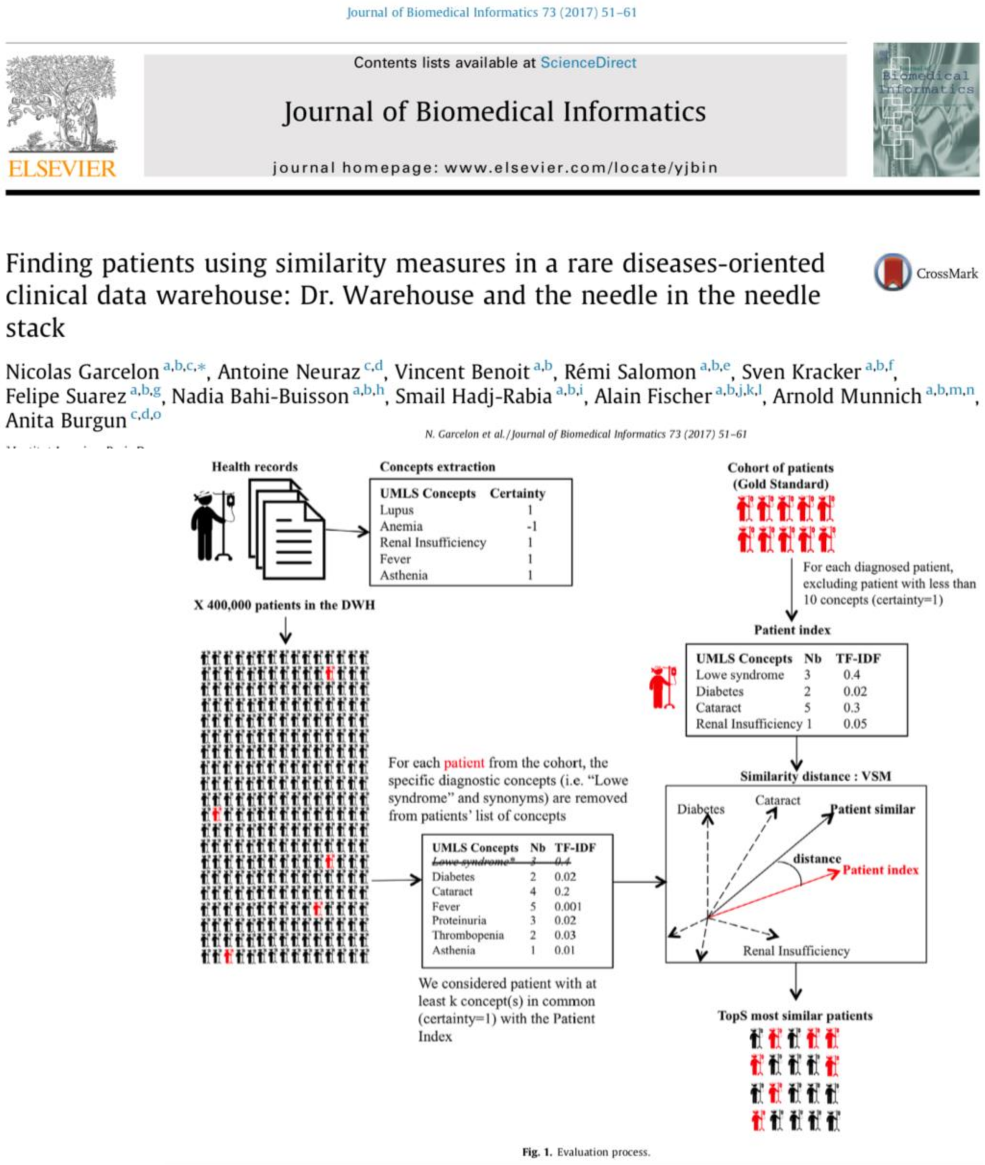
« we made the decision on the basis of the best data available »
« in the light of experience as guided by intelligence. »



Garcelon N. et al. J. Biomed Inf 2017
Chen X. et al. J. Biomed Inf. 2019

RESPONSIBLE INNOVATION

Can decision be automated through machine learning?


Journal of Biomedical Informatics 73 (2017) 51–61

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

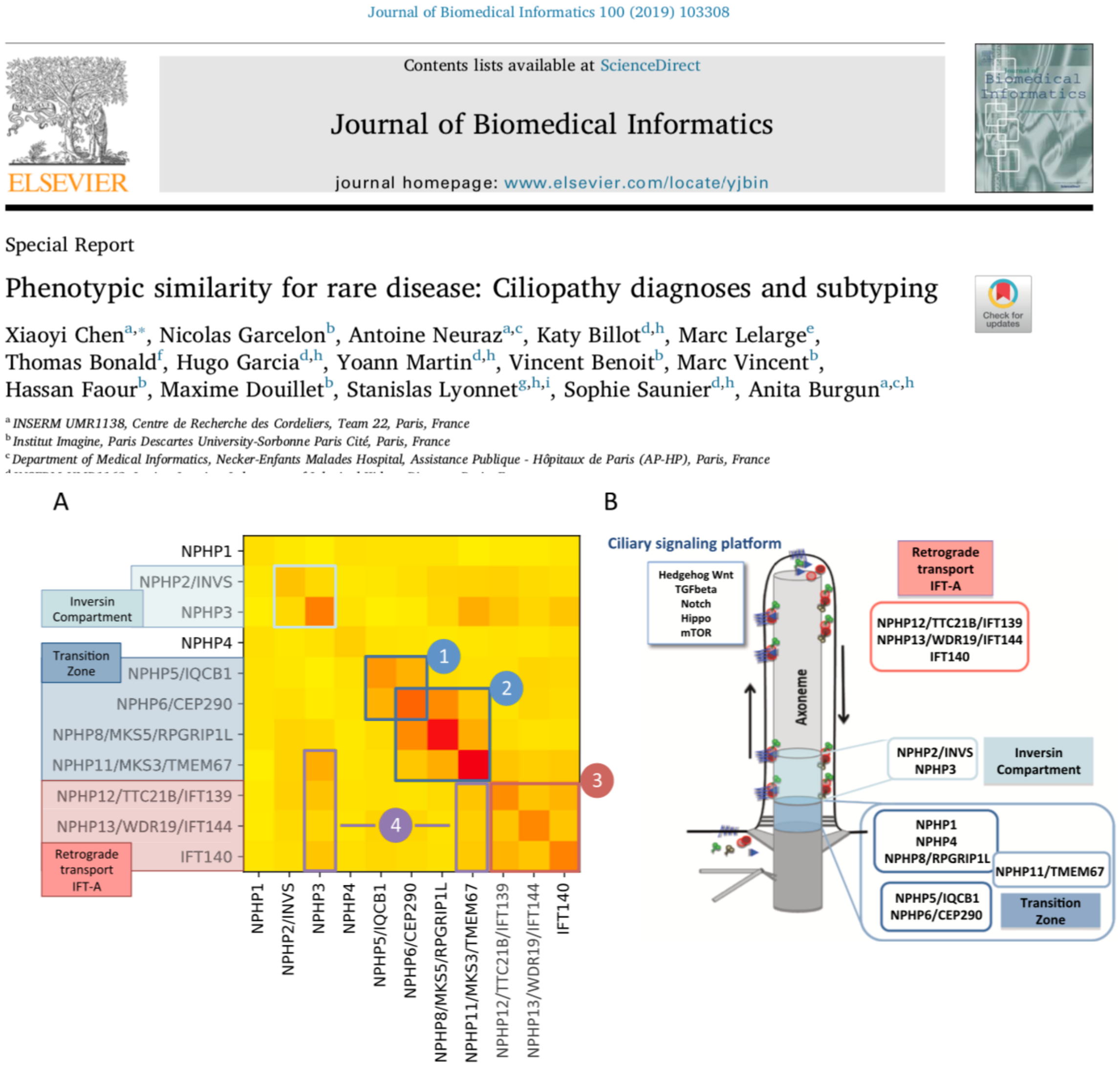
ELSEVIER

Finding patients using similarity measures in a rare diseases-oriented clinical data warehouse: Dr. Warehouse and the needle in the needle stack

Nicolas Garcelon [a,b,c,*], Antoine Neuraz [c,d], Vincent Benoit [a,b], Rémi Salomon [a,b,e], Sven Kracker [a,b,f], Felipe Suarez [a,b,g], Nadia Bahi-Buisson [a,b,h], Smail Hadj-Rabia [a,b,i], Alain Fischer [a,b,j,k,l], Arnold Munnich [a,b,m,n], Anita Burgun [c,d,o]

N. Garcelon et al. / Journal of Biomedical Informatics 73 (2017) 51–61

Fig. 1. Evaluation process.


Journal of Biomedical Informatics 100 (2019) 103308

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

ELSEVIER

Special Report

Phenotypic similarity for rare disease: Ciliopathy diagnoses and subtyping

Xiaoyi Chen[a,*], Nicolas Garcelon[b], Antoine Neuraz[a,c], Katy Billot[d,h], Marc Lelarge[e], Thomas Bonald[f], Hugo Garcia[d,h], Yoann Martin[d,h], Vincent Benoit[b], Marc Vincent[b], Hassan Faour[b], Maxime Douillet[b], Stanislas Lyonnet[g,h,i], Sophie Saunier[d,h], Anita Burgun[a,c,h]

[a] INSERM UMR1138, Centre de Recherche des Cordeliers, Team 22, Paris, France
[b] Institut Imagine, Paris Descartes University-Sorbonne Paris Cité, Paris, France
[c] Department of Medical Informatics, Necker-Enfants Malades Hospital, Assistance Publique - Hôpitaux de Paris (AP-HP), Paris, France

Genetics inMedicine

www.nature.com/gim

Check for updates

BRIEF COMMUNICATION

Deep phenotyping unstructured data mining in an extensive pediatric database to unravel a common *KCNA2* variant in neurodevelopmental syndromes

Marie Hully[1], Tommaso Lo Barco[1], Anna Kaminska[1,2], Giulia Barcia[1,3], Claude Cances[4], Cyril Mignot[5], Isabelle Desguerre[1], Nicolas Garcelon[6,7], Edor Kabashi[8] and Rima Nabbout[1,8]✉

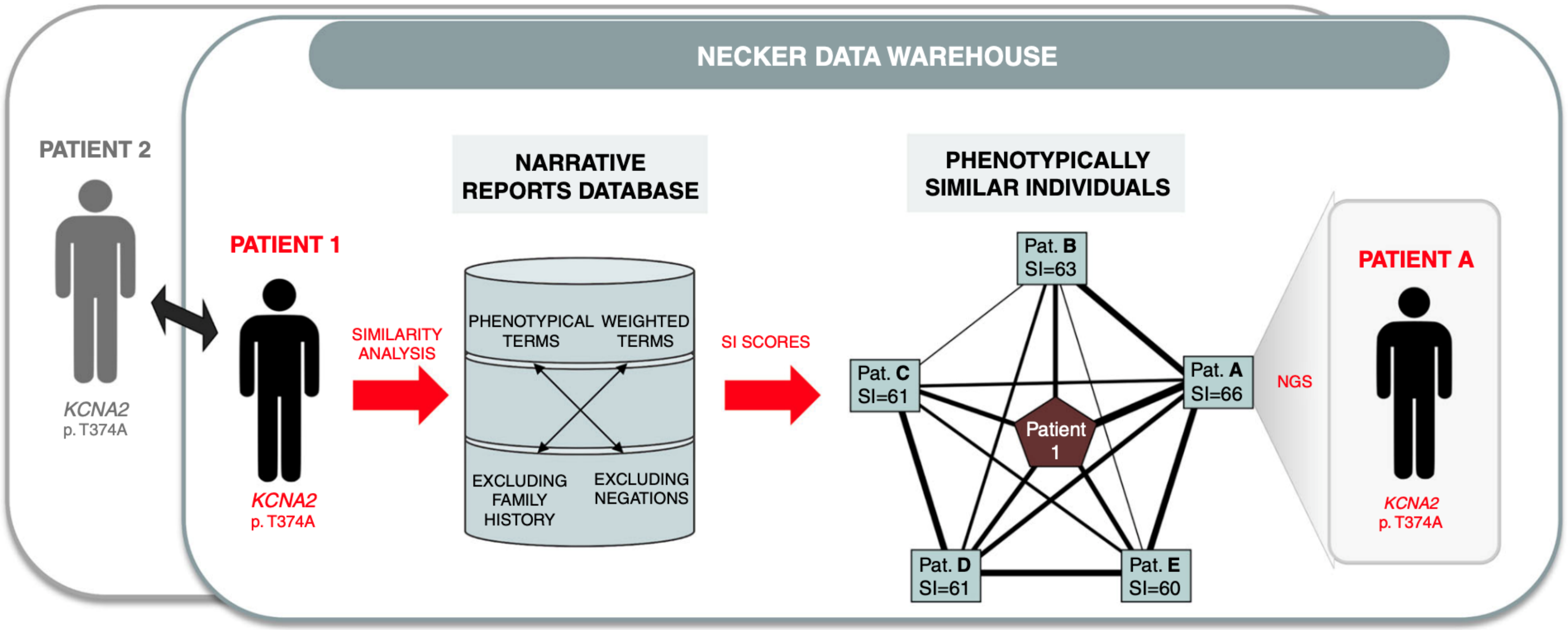


Fig. 1 Display of the two patients (patient 1 from our institution and patient 2 from another institution in our reference center network) sharing the same phenotype and the same *KCN2A* variant. Similarity analysis with all data warehouse narrative reports was performed, yielding a high similarity index (SI) in five patients (patients A–E). Exome sequencing validated that patient A, who had the highest SI, harbored the same *KCNA2* variant. NGS next-generation sequencing.

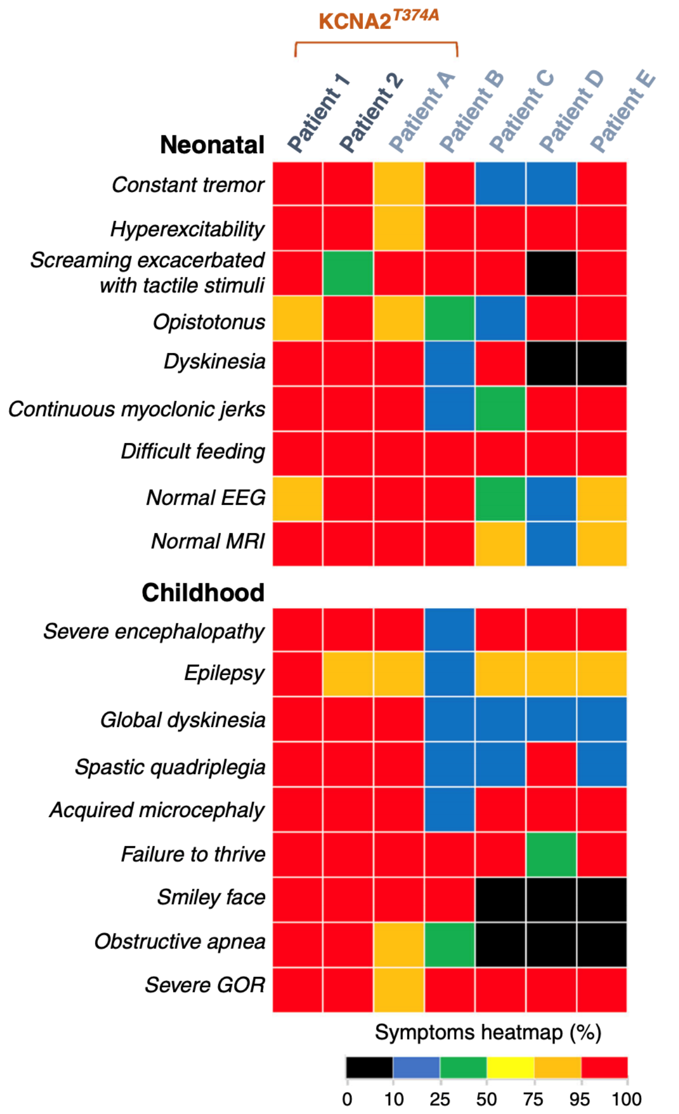


Fig. 2 Clinical heat map describing the detailed characteristics of the patients in this study. Heatmap for patient 1 and 2 with

Can decision be automated through machine learning?

- International guidelines for HBP control
- 17856 patients /1st visit
- Women are prescribed more diuretics and beta-blokers

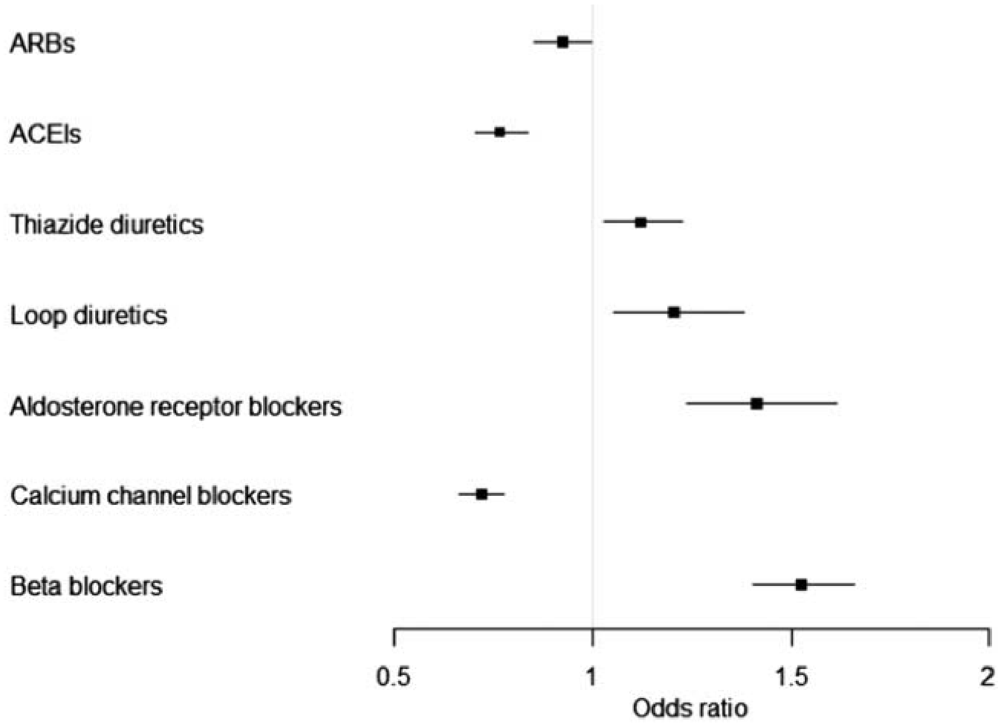


FIGURE 1 Odds of women being treated by a given class of antihypertensive drug at the first consultation compared with men, after adjusting for all available potential confounding variables.

Deborde et al. Sex differences in antihypertensive treatment in France among 17 856 patients in a tertiary hypertension unit. J Hypertens. 2018

Can decision be automated through machine learning?

TYRANNIZED BY THE DATA

- International guidelines for HBP control
- 17856 patients /1st visit
- Women are prescribed more diuretics and beta-blokers

- **Guidelines vs machine learning would lead to different drug prescription**

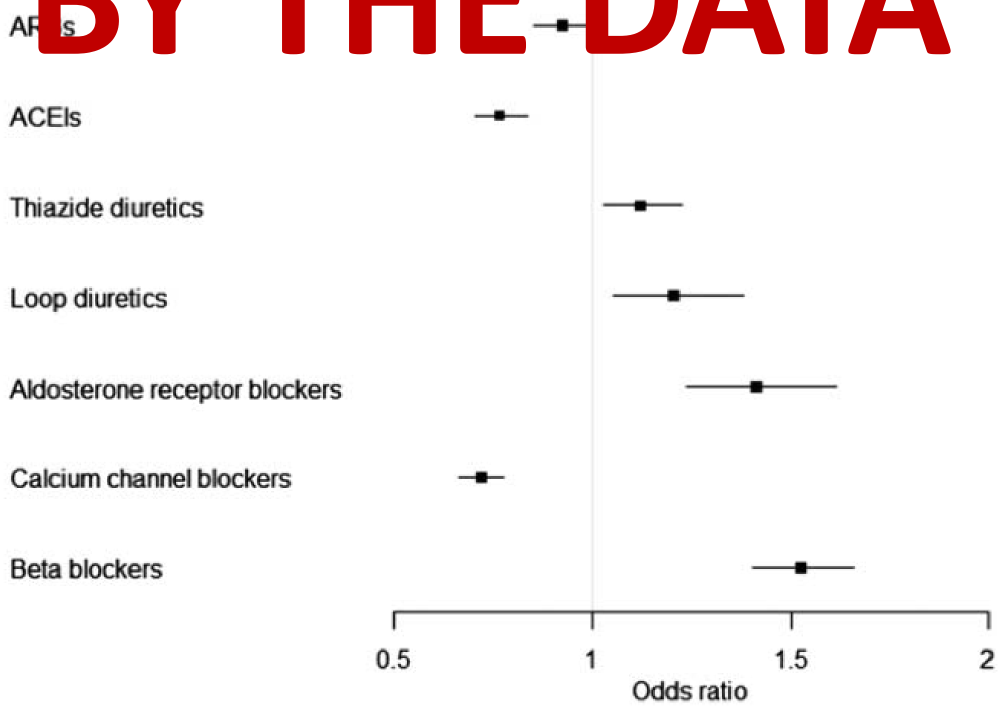


FIGURE 1 Odds of women being treated by a given class of antihypertensive drug at the first consultation compared with men, after adjusting for all available potential confounding variables.

Deborde et al. Sex differences in antihypertensive treatment in France among 17 856 patients in a tertiary hypertension unit. J Hypertens. 2018

Can decision be automated through machine learning?

TYRANNIZED BY THE PAST

- nter ati na g ide ir es r HBP control
- 17856 patients /1st visit
- Women are prescribed more diuretics and beta-blokers

- **Guidelines vs machine learning would lead to different drug prescription**

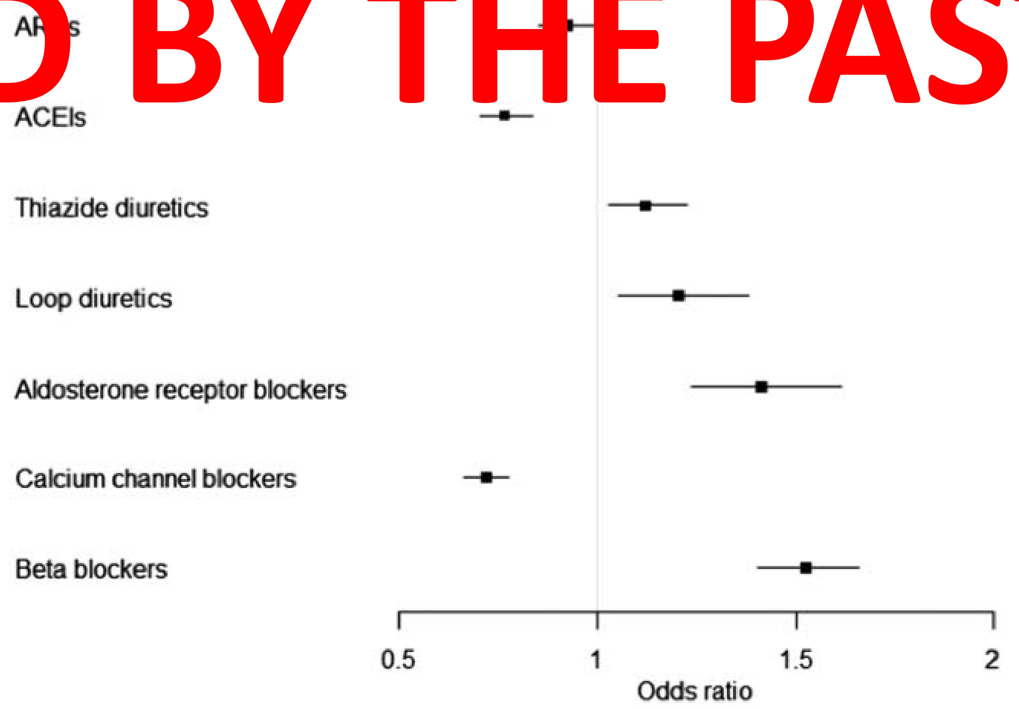


FIGURE 1 Odds of women being treated by a given class of antihypertensive drug at the first consultation compared with men, after adjusting for all available potential confounding variables.

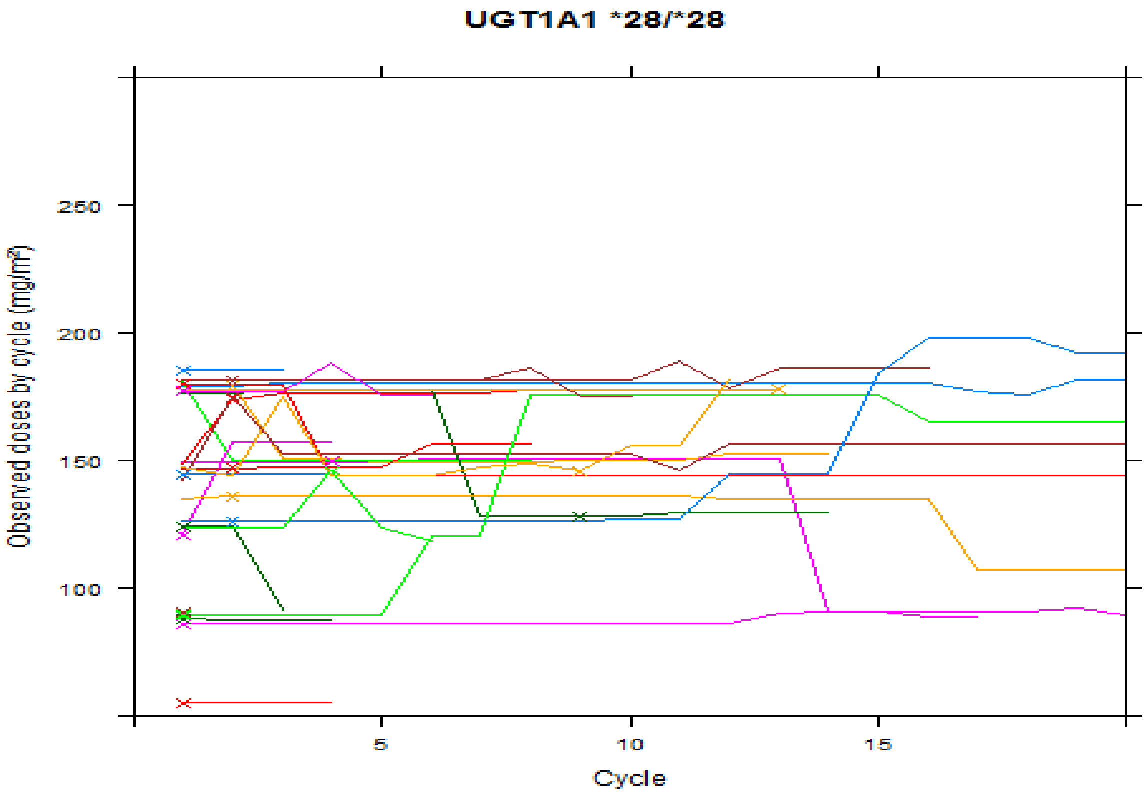
Deborde et al. Sex differences in antihypertensive treatment in France among 17 856 patients in a tertiary hypertension unit. J Hypertens. 2018

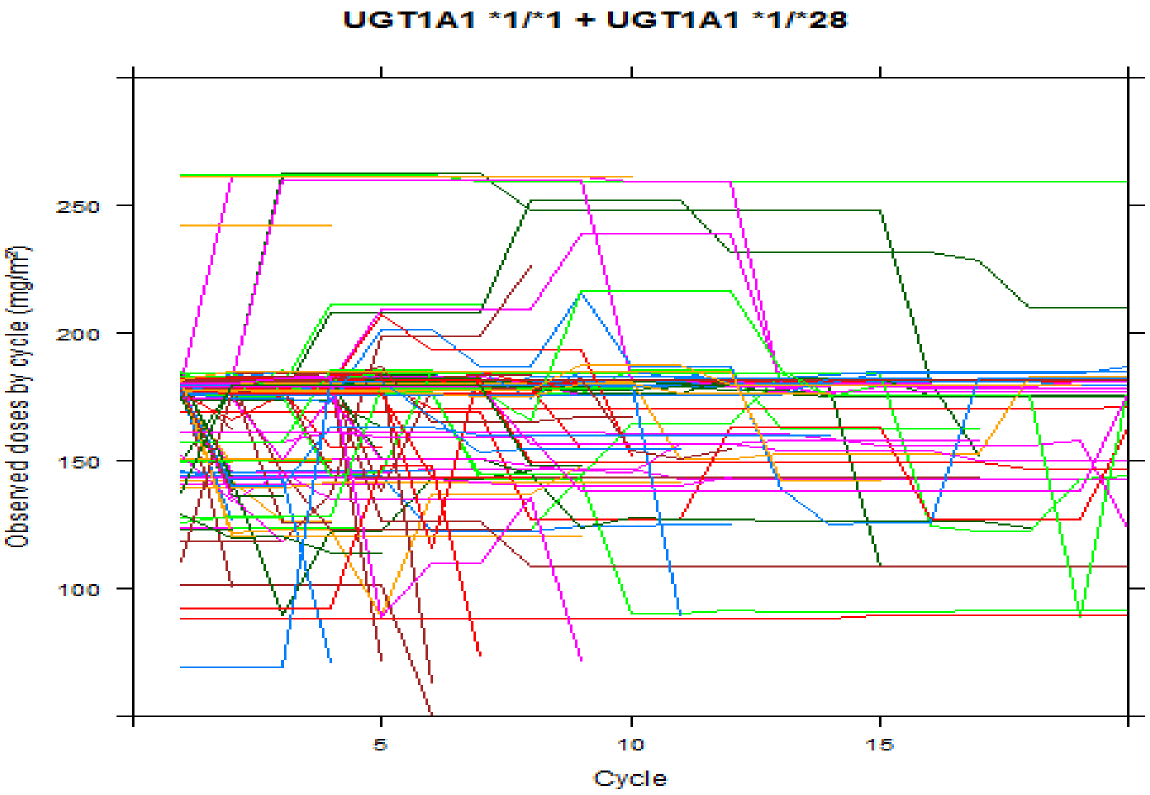
La Garantie Humaine dans le projet de règlement sur l'IA de la Commission européenne !

Ethik-IA, JEUDI 22 AVRIL 2021 Soyez le premier à réagir

Le principe de Garantie Humaine de l'IA (Human Oversight) est introduit à l'article 14 du projet de règlement sur l'intelligence artificielle diffusé ce jour par la Commission européenne. Ce faisant, l'article 14 donne une portée applicative générale pour l'ensemble des champs et secteurs d'usage de l'IA à ce principe proposé en 2017 par Ethik-IA dans le domaine de la santé et qui n'a cessé, depuis lors, de faire l'objet de reconnaissances de plus en plus larges. Ce principe s'inscrit dans une logique de régulation positive visant à soutenir le développement de l'intelligence artificielle en France et en Europe, dans un cadre permettant d'en réguler les risques éthiques.

Dose profiles





Decision making and explainability (Boulet SMMR 2019)

Table 2 Clinical relevance weights for each covariate elicited from each physician.

	Physicians											
Variables	1				2				3			
Age $\geq$ 80 years	100	-	-	-	60	-	-	-	100	-	-	-
Weight loss $>$ 10%	50	-	-	-	20	-	-	-	50	-	-	-
WHO score (1,2,3,4)	0	20	20	20	0	0	40	100	0	0	80	100
Bilirubin $>$ 35, $>$ 50 μmol/L	100	100	-	-	40	80	-	-	100	100	-	-
Treatment line 3, $>$ 3	30	50	-	-	0	0	-	-	0	0	-	-
Toxicity grades 1, 2, 3, 4												
Vomiting	0	20	80	90	0	30	70	100	0	10	10	10
Nausea	0	20	80	-	0	10	50	-	0	10	10	-
Diarrhea	0	40	80	100	0	20	50	100	0	50	80	100
Asthenia	10	50	100	-	10	10	40	-	0	0	70	-
Neutropenia	0	70	100	100	0	0	30	50	0	0	50	50
Thrombopenia	40	100	100	100	0	0	20	30	0	0	50	50
Anemia	0	50	80	100	0	0	20	30	0	0	0	0

	Physicians											
Variables	4				5				6			
Age $\geq$ 80 years	80	-	-	-	100	-	-	-	100	-	-	-
Weight loss $>$ 10%	80	-	-	-	50	-	-	-	60	-	-	-
WHO score (1,2,3,4)	0	20	80	100	0	30	100	100	0	70	100	100
Bilirubin $>$ 35, $>$ 50 μmol/L	20	80	-	-	100	100	-	-	100	100	-	-
Treatment line 3, $>$ 3	0	0	-	-	0	0	-	-	0	50	-	-
Toxicity grades 1, 2, 3, 4												
Vomiting	10	20	80	100	0	30	70	100	0	0	70	100
Nausea	10	30	80	-	0	10	40	-	0	0	30	-
Diarrhea	0	20	70	90	0	20	50	100	0	50	100	100
Asthenia	10	50	70	-	0	20	50	-	0	40	80	-
Neutropenia	0	20	70	80	0	0	20	100	0	0	0	50
Thrombopenia	0	50	80	100	0	0	20	70	0	0	0	0
Anemia	0	20	50	70	0	0	30	80	0	0	0	0

Lyell *et al. BMC Medical Informatics and Decision Making* (2017) 17:28
DOI 10.1186/s12911-017-0425-5

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

Automation bias in electronic prescribing

David Lyell[1*], Farah Magrabi[1], Magdalena Z. Raban[2], L.G. Pont[2], Melissa T. Baysari[2,3], Richard O. Day[4] and Enrico Coiera[1]

Quand le CDSS est correct

REDUIT LES ERREURS DE 58,8%

Quand le CDSS se trompe

AUGMENTE LES ERREURS DE 86,6%

Trust the CDSS?

CDSS that make less errors

Supervision by humans

Decision in medicine

- Decision-making is a key activity for doctors

- Haynes *et al*

‘It is a guide for thinking about how decisions should be made rather than a schema for how they are made’.

- Russ Altman, Stanford University Institute for Human-Centered Artificial Intelligence

“As AI starts to impact all areas of medical discovery and healthcare delivery, the focus should be how it improves care, leading to longer and happier lives”

anita.burgun@aphp.fr