



Identification des lignes de traitement en cancérologie à partir des données du SNDS :

Une méthode d'intelligence artificielle basée sur
l'alignement de séquences

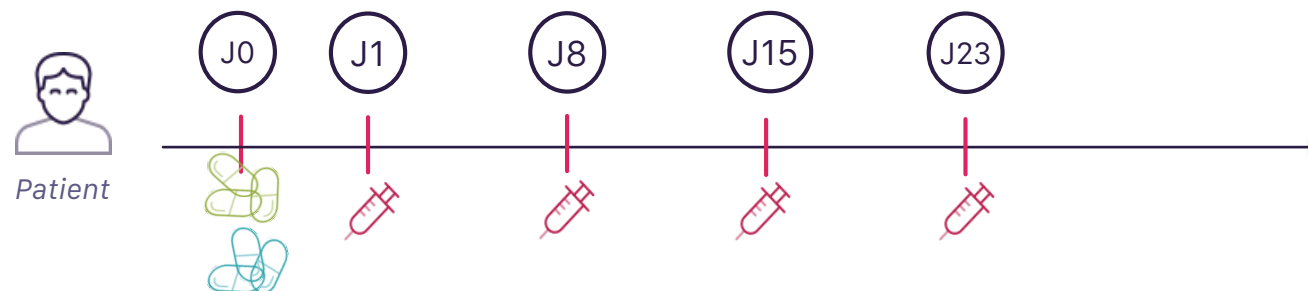
Martin PRODEL, Fanny RAGUIDEAU, Ludovic LAMARSALLE

EMOIS 2019, Nancy

Objectifs

Utiliser les données du SNIIRAM pour identifier les prises en charge des patients en cancérologie.

HEVA



- Identifier les **protocoles de soins théoriques** suivis au sein des données
- Estimer le nombre de patients suivant un niveau de **lignes de traitements**

Utilisation finale :

- Déterminer la réalité des prises en charge sur tout le pays
- Identifier d'éventuelles déviations ou pistes d'amélioration
- Apporter une connaissance sur l'épidémiologie en cancérologie

Contraintes & Défis



Contraintes liées aux données

- a) Médicaments **non observables** (+++ dans SNIIRAM)
- b) **Variations** entre recommandations cliniques et données observées
- c) **Durée de suivi importante** (jusque 10 ans) → complexité des données, présence de bruit et qualité du recueil

Défis opérationnels

- a) Une méthode automatisée → milliers de patients (vs. validation manuelle par médecin)
- b) Quantifier la vraisemblance entre 1 patient et le protocole théorique (vs. règles oui/non)
- c) Une méthode flexible dans les comparaison : démarche « I.A. » (vs. règles strictes)



La méthode : A.T.L.A.S.

by HEVA

ATLAS : Analysis of Treatment Lines
using Alignment of Sequences

HEVA

A l'origine

Comparer 2 séquences d'ADN



Séquence 1 T G T T A C G G

Séquence 2 G G T T G A C T A

Quelles sections sont
les plus semblables ?

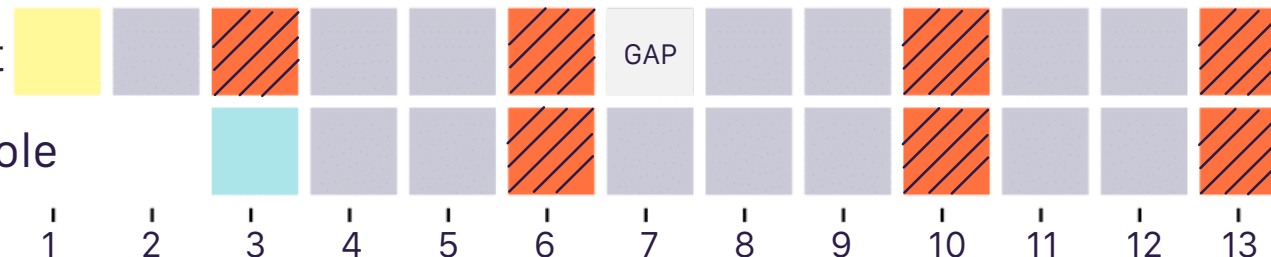
Maintenant

Aligner un patient vs. un protocole

Score : 98% (a gap makes a 2% loss)

Patient

Protocole



Légende



La méthode ATLAS

L'ALGORITHME D'ALIGNEMENT EN 3 ETAPES

HEVA

1. PARAMETRAGE INITIAL

PROTOCOL PATIENT

	Pause	A _{hosp.}	B _{hosp.}	C _{pharmacy}	B+C
Pause	0.5	-1	-1	-1	-1
A	-1	5	2	-5	1
B	-1	2	5	-2	$X_{C,B}$
C	-1	-5	-2	5	$X_{B,C}$
B+C	-1	1	$X_{C,B}$	$X_{B,C}$	10

$X_{(u,v)} = \begin{cases} 10 & \text{if } u \text{ is found within 3 days of } v, \text{ thus matching "u + v"} \\ 4 & \text{otherwise} \end{cases}$

G_v

	Pause	Drug	Gap
Pause	$-1/2$	-1	$-1/2$
Drug	-1	-2	-1
Gap	$-1/2$	-1	$-1/4$

G_h

	Pause	Drug	Gap
Pause	-1	-3	-1
Drug	-3	-5	-2
Gap	-1	-2	-1

2. POUR 1 PATIENT vs. 1 PROTOCOLE

	Seq2	Drug B	Pause	Pause	Drug C	Pause	Drug B	Pause	Pause
Seq1	0	0	0	0	0	0	0	0	0
Drug A	0	0	3	1	0	0	0	3	3
Pause	0	0	3	1	0	0	0	3	6
Pause	0	3	1	6	4	2	0	1	4
Pause	0	3	1	4	9	7	5	3	2
Drug A	0	1	6	4	7	6	4	8	6
Pause	0	0	4	3	5	10	8	6	5
Drug B	0	0	2	1	3	8	13	11	9
Drug C	0	3	1	5	4	6	11	10	8
Pause	0	1	0	3	2	7	9	8	7

3. RESULTATS POUR 1 PATIENT, SUR CHAQUE SEGMENT

Classement des alignements :

1. Avec protocole 8 : 95%
2. Avec protocole 2 : 75%
3. Avec protocole 1 : 40%
- ...



Possibilité de déterminer toute l'histoire du patient

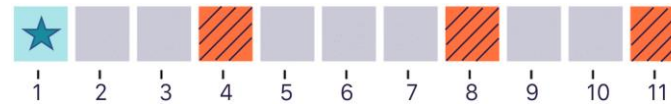
P8 → P8 → P8 → P8 → P1 → P1
=
Ligne 1 (P8) → Ligne 2 (P1)

Expérimentation générique

15 protocoles types

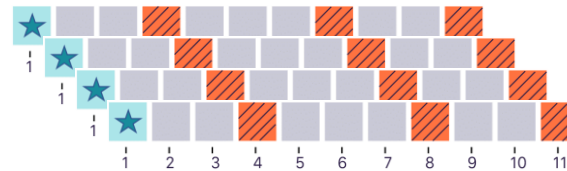
ID	Drug scheme (Name: administration days, sequence <>)
P1	A: D1-D4, B: D1-D4 < (AB)(AB)(AB)(AB) >
P2	A: D1-D4, B: D1-D4, C: D1-D28, < (ABC)(ABC)(ABC)(ABC)CCC...C >
P3	A: D1-D4, B: D1-D4, D: D1/D8/D15/D22 < (ABD)(AB)(AB)(AB) ppp D pppppp D ppppppp D >
P4	D: D1/D8/D15/D22, X: D1/D8/D15, Y: D1-D8 < (DXY)YYYYYY(DXY) pppppp (XD) pppppp D >
P5	E, F, Y: D1-D4 < (EFY)(EFY)(EFY)(EFY) >
P6	D, Y: D1/D4/D8/D11 < (DY) pp (DY) ppp (DY) pp (DY) >
P7	D: D1/D4/D8/D11, C: D1-D21, Y: D1-D4/D8-D11 < (DCY)(CY)(CY)(DCY)CCCC(DCY)(CY)(CY)(DCY) CCCCCCCCCC >
P8	D: D1/D4/D8/D11, C: D1-D28, I: D1/D8/D15/D22 Y: D1-D2/D8-D9/D15-16/D22-23, < (DCIY)(CY)CDCCC(DIY)(CY)CDCCC(CIY)(CY) CCCCC(CY)(CY)CCCC >
P9	D: D1/D4/D8/D11, H: D1-D4, Y: D1-D2/D4-D5/D8-D9/D11-12 < (DHY)(HY)H(DHY)Y pp (DY)Y p (DY)Y >
P10	D: D1/D4/D8/D11, F: D1-D4, Y: D1-D4/D8-D11/D15-D18 < (DFY)(FY)(FY)(DFY) ppp (DY)YY(DY) ppp YYY >
P11	G: D1/D3/D5/D8/D10/D12, D, Y: D1/D4/D8/D11 < (GDY) p G(DY)G pp (GDY) p G(DY)G >
P12	J: D1-D2/D8-D9/D15-16, Y: D1/D8/D15, H: D1-D22 < (JYH)(JH)HHHHH(JYH)(JH)HHHHH (JYH)(JH)HHHHH >
P13	L: D1-D2/D8-D9/D15-16, Y: D1/D8/D15, H: D1-D22 < (LYH)(LH)HHHHH(LYH)(LH)HHHHH (LYH)(LH)HHHHH >
P14	Y: D1/D8/D15, H: D1-D22 < (YH)HHHHHH(YH)HHHHHH(YH)HHHHHH >
P15	K: D1/D8, V, Y: D1/D8/D15/D22 < (KVY) pppppp (KVY) pppppp (VY) pppppp (VY) >

Pour chaque protocole X



Création de 1000 séquences

- Exactement identiques au protocole X
- Jouant le rôle du patient



...

Alignement de chaque patient

HEVA

- Avec chacun des 15 protocoles
- Avec l'objectif de retrouver le protocole X comme le meilleur
- Evaluation du taux de succès

Bruitage des séquences patients

- Perturbations dans la séquence
- Ajout, retrait, décalage



Expérimentation : les résultats

HEVA

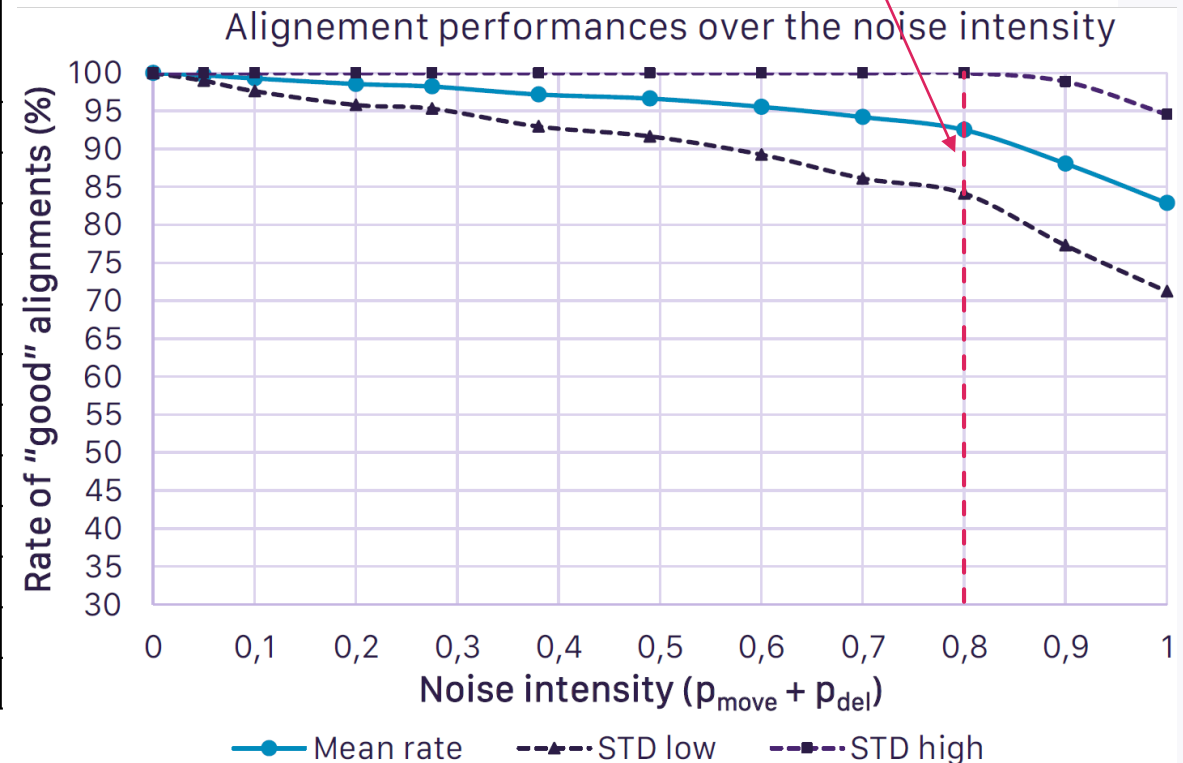
Bruit faible

Noise settings		Good			Bad		
p_{move}	p_{del}	%	STD	N	%	STD	N
0	0	100.0	± 0	1000	0.0	± 0.0	0
0.05	0.0	99.7	± 0.7	997.1	0.3	± 0.7	2.9
0.1	0.0	99.3	± 1.7	993.0	0.7	± 1.7	7.0
0.15	0.05	98.6	± 2.8	985.7	1.4	± 2.8	14.2
0.2	0.075	98.2	± 2.9	982.3	1.8	± 2.9	17.6
0.3	0.08	97.2	± 4.2	971.9	2.8	± 4.2	27.7
0.4	0.09	96.6	± 5.0	966.5	3.2	± 4.9	32.1
0.5	0.1	95.6	± 6.3	955.6	4.2	± 6.1	41.5
0.55	0.15	94.2	± 8.1	942.1	5.2	± 7.6	51.8
0.6	0.2	92.5	± 8.4	925.2	6.2	± 8.0	62.5
0.6	0.3	88.1	± 10.8	880.9	9.0	± 10.7	90.2
0.6	0.4	82.9	± 11.6	828.9	10.9	± 11.0	109.0

Bruit très élevé

Niveau de « bruit »
attendu du SNIIRAM¹

Seuil de non
utilisation d'ATLAS
(recommandation)



1. A. Palmaro, M. Gauthier, F. Despas, and M. Lapeyre-Mestre, "Identifying cancer drug regimens in french health insurance database: An application in multiple myeloma patients," *Pharmacoepidemiology and Drug Safety*, vol. 26, 07 2017.

Publier

Une méthode robuste et performante pour identifier les lignes de traitements dans le SNIIRAM : **publication dans une revue en cours**.

Eprouver

Mise en œuvre dans une aire thérapeutique précise.

Affiner

- Une méthode statistique générique pour le calcul des paramètres (matrices)
- Mesurer la similarité entre protocoles théoriques pour ajuster ATLAS
- Extrapoler des connaissances sur un échantillon vers la cohorte entière



Merci pour votre attention

Martin : mprodel@hevaweb.com