Prédiction basée sur l'apprentissage automatique du statut mutationnel *EGFR* dans les cancers broncho-pulmonaires non à petites cellules

Matthieu Guillard, Arnaud Uguen, Marion Le Rochais, Pierre Le Noac'h, Rachid Zeghlache, Ikram Brahim

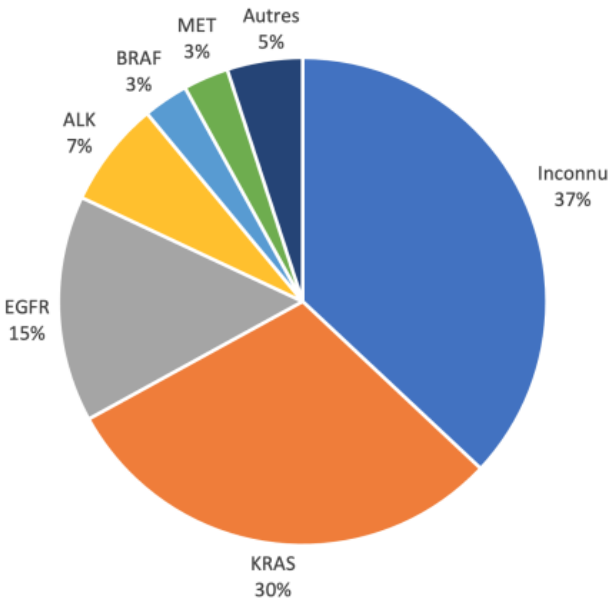
UMR1227 LBAI Brest

Introduction

Adénocarcinome pulmonaire

- Présence d'un driver oncogénique dans certains cas
  → EGFR, ALK

- Accessible à un traitement ciblé dès la première ligne

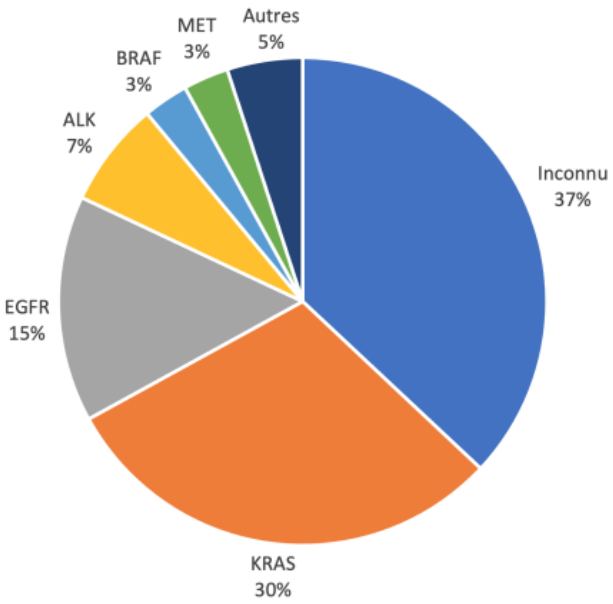


Introduction

Mutations recherchées en biologie moléculaire

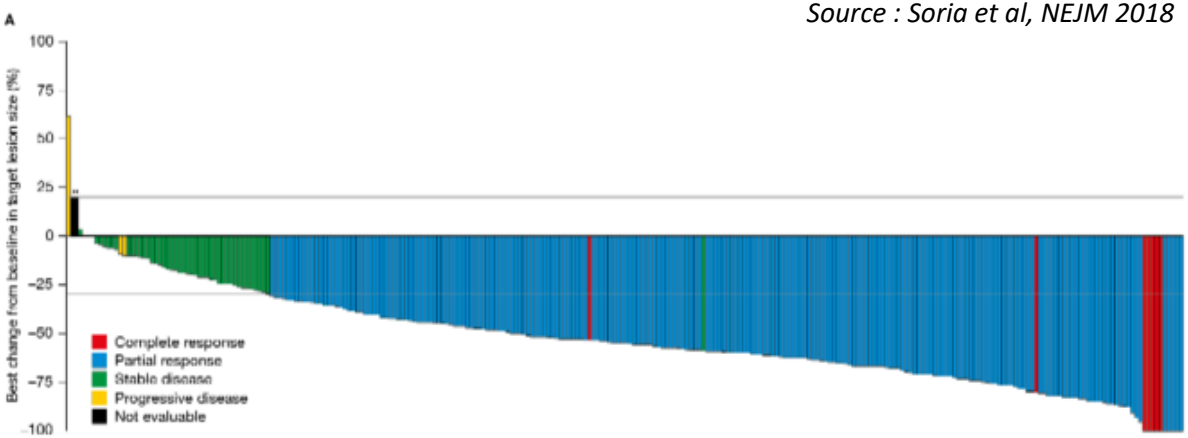
- Avantages : gold standard, multiples altérations détectables

- Inconvénient : coût, délai de résultat

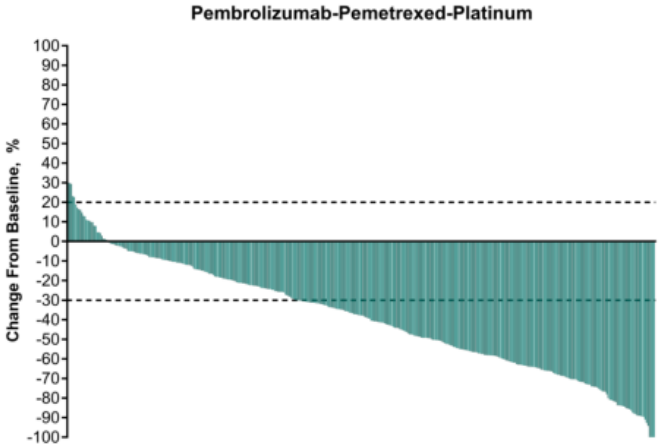


Introduction

Source : Soria et al, NEJM 2018



Taux de réponse aux TKI : 80%

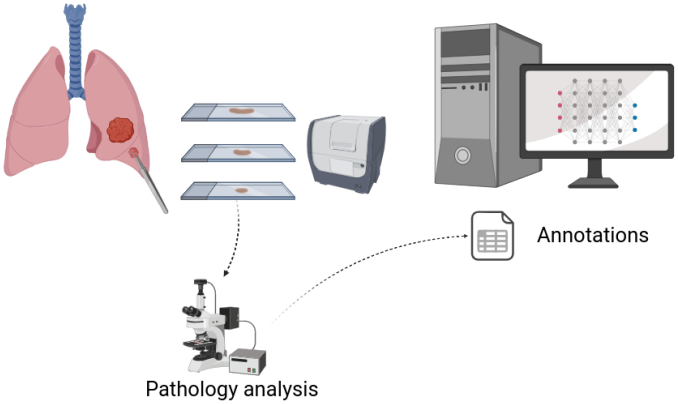


Taux de réponse CT + IO : 45%

Source : Gandhi et al, NEJM 2018

Objectif

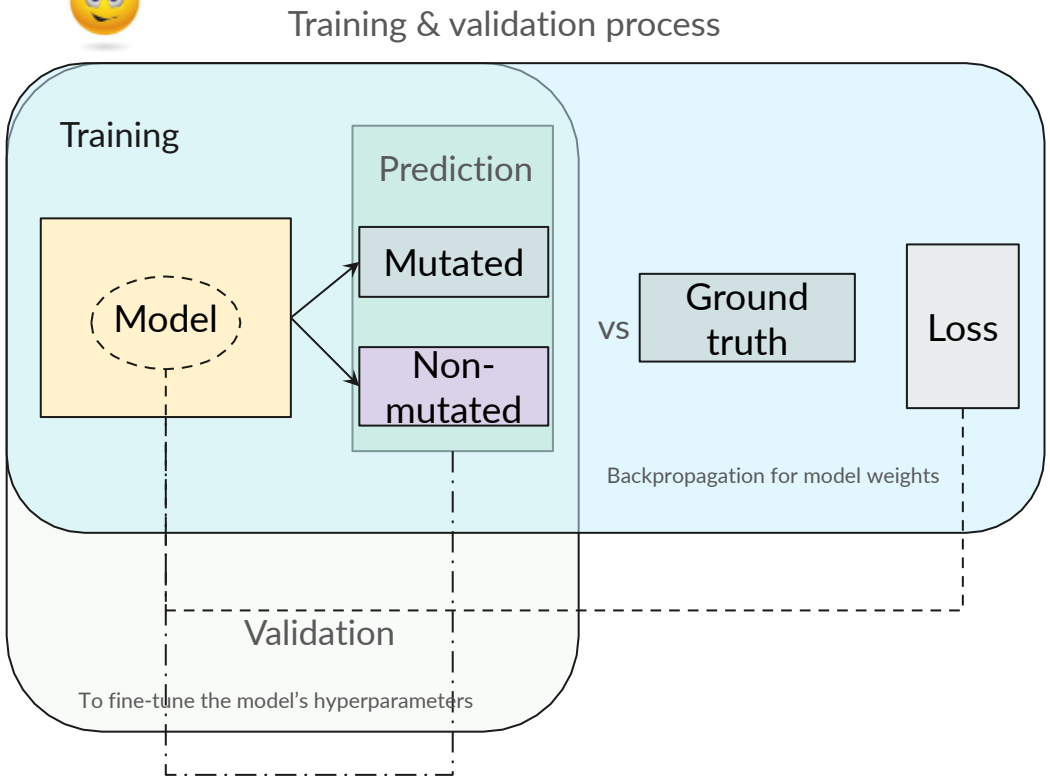
Est-il possible de déterminer le statut *EGFR* à partir d'une lame HES à l'aide de l'IA ?



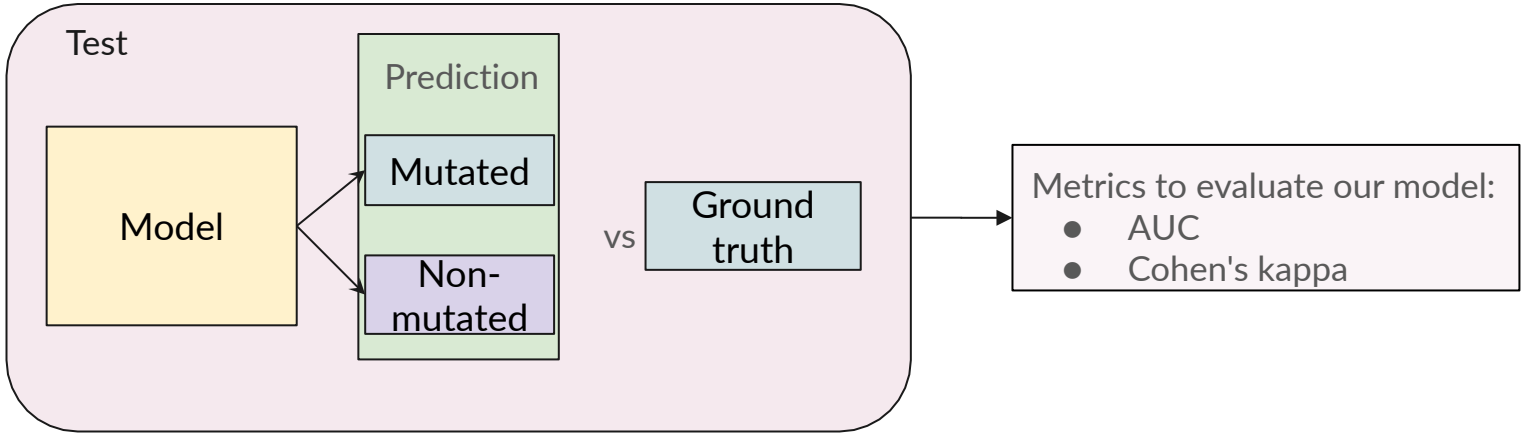
Préparation des données

Répartition de 24 cas :

- Set d'entrainement : 10 cas
  → 5 mutés et 5 wild type
- Set de validation : 9 cas
  → 4 mutés et 5 wild type
- Set de test : 5 cas
  → 3 mutés et 2 wild type

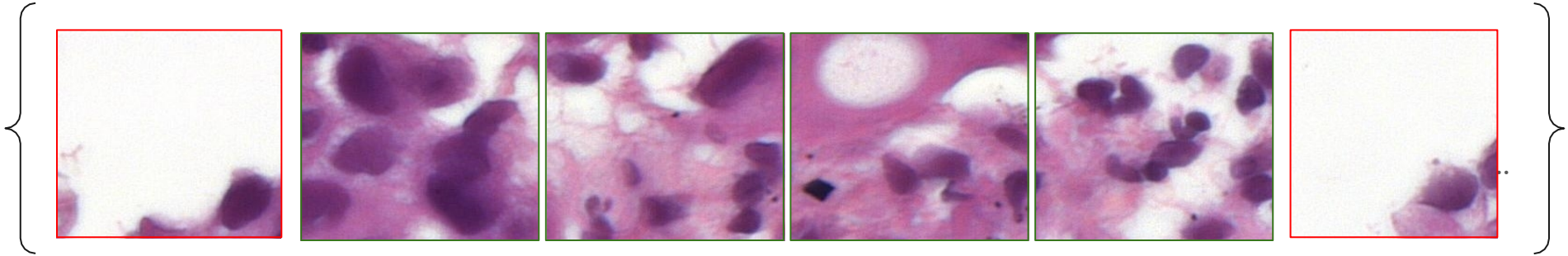


Préparation des données



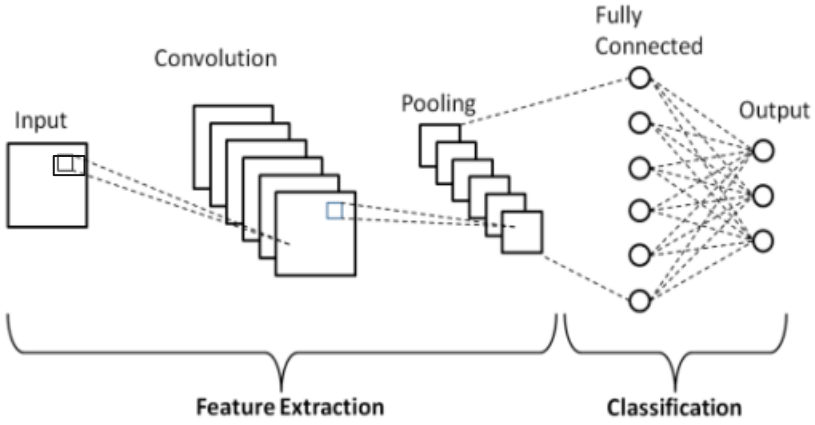
Préparation des données

- Annotation de lames entières : muté / wild type
- Découpage en tuiles et élimination des zones a-tissulaires
- En moyenne 3000 tuiles par lame : 500 choisies au hasard par cas pour entraîner le modèle



Modèle

1. Input Layer :
   a. Receives input data tiles
2. Convolutional Layers :
   a. Local features detected using convolutional operations.
   b. Filters are applied to input images.
3. Pooling Layers :
   a. Reduces spatial dimensions of the input volume.
4. Fully Connected Layers :
   a. Neurons are connected to all neurons in the previous layer.
5. Outputs a probability indicating the likelihood of each class.



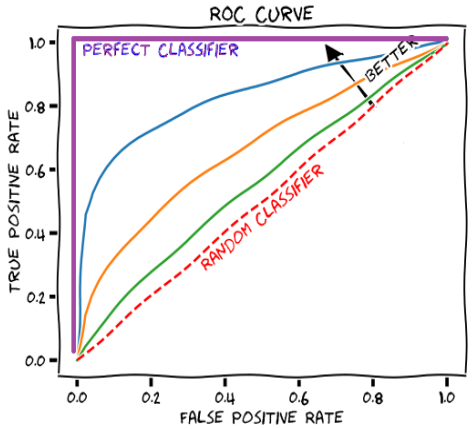
ikrammeki@gmail.com

Résultats

- AUC : 0,83

  Ratio du taux de vrai positifs vs taux de faux positifs
- Kappa de Cohen : 0.62.

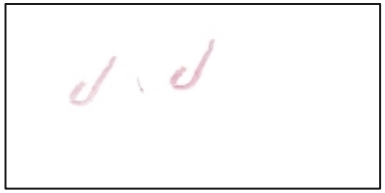
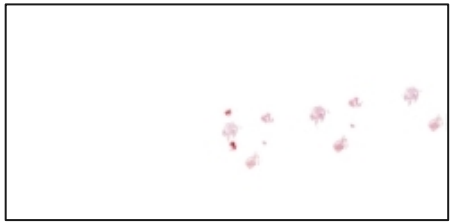
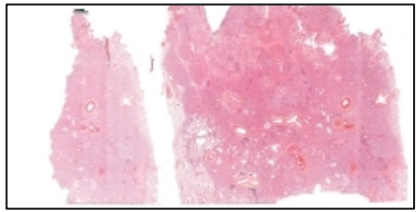
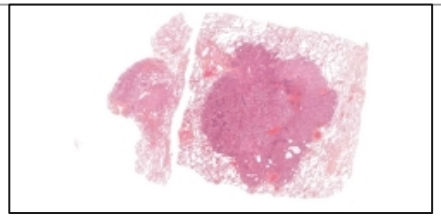
  Accord entre l'annotateur et la prédiction du modèle

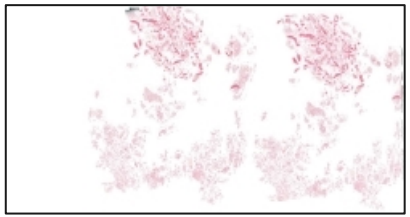


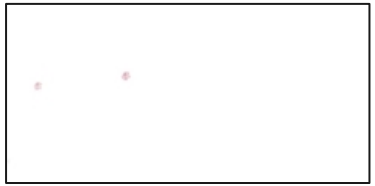
Cohen's Kappa	Interpretation
0	No agreement
0.10 - 0.20	Slight agreement
0.21 - 0.40	Fair agreement
0.41 - 0.60	Moderate agreement
0.61 - 0.80	Substantial agreement
0.81 - 0.99	Near perfect agreement
1	Perfect agreement

		Actual Outcome	
		Positive	**Negative**
Predicted Outcome	**Positive**	True Positive	False Positive
	Negative	False Negative	True Negative

…mais dans une nouvelle série de la “vraie vie”…

Patient	Probability prediction	Ground truth	Biopsy
E1	0.72	*EGFR* del exon 19	
E2	0.30	*EGFR L858R*	
E3	0.46	*EGFR* del exon 19	
E4	0.29	*EGFR* del exon 19	

Patient	Probability prediction	Ground truth	Biopsy
E5	0.35	*EGFR WT* *KRASG12V*	
E6	0.97	*EGFR S768I & G719C*	
E7	0.80	*EGFR WT* *KRASG61L*	
E8	0.30	*EGFR WT* *KRASG12V*	

Patient	Probability prediction	Ground truth	Biopsy
E9	0.75	*EGFR WT* *KRASG12V*	
E10	0.70	*EGFR WT* *KRASG12C*	
E11	0.81	*EGFR WT* *KRASG12C*	
E12	0.54	*EGFR* del exon 19	

…mais dans une nouvelle série de la “vraie vie”…

Discussion

- Quelle serait l'utilité d'un tel modèle en pratique courante ?
  → Outil de screening : permettrait d'orienter directement vers le test de la mutation recherchée (test ciblé rapide vs test panel de gène) avec rendu + rapide, coût moins élevé

- Pourrait remplacer le test de biologie moléculaire ?
  → Avis de l'oncologue : que la p soit < 0.1 ou p > 0.95 = pas de traitement sans "100%" de résultat de biologie moléculaire! = +++ faire le test peu importe la prédiction!

Conclusion

- Les analyses de “pathomique” constituent un enjeu de la pathologie et de l’oncologie de demain…
- … mais à condition de dépasser le “faisable” pour se concentrer sur le “vraiment utile”
  = peu (à pas) d’intérêt à prédire des statuts mutationnels accessibles via des techniques de pathologie moléculaire en pratique…