



Sciences Economiques & Sociales de la Santé
& Traitement de l'Information Médicale

sesstim.univ-amu.fr

Raphaël PORCHER

MCU - PH

Centre de Recherche Épidémiologie et Statistique (CRESS) Sorbonne Paris Cité, France
UMR1153 Université Paris Descartes / INSERM

Méthodes d'identification de patients répondeurs à un traitement

janvier 2017



Cliquez ici pour voir l'intégralité des ressources associées à ce document

Webinar QuanTIM
20 janvier 2017

Méthodes d'identification des patients répondeurs à un traitement

Raphaël Porcher

INSERM U1153, Université Paris Descartes, AP-HP, France



Médecine personnalisée, stratifiée, de précision

- Termes différents, même concept
- Objectif : « donner le bon médicament au bon patient »
- Fortement lié aux avancées en génétique et autres analyses moléculaires et cellulaire
- Terme utilisé pour la première fois pour la génétique
- Mais le concept est plus large

Attractif

- Concept qui a attiré l'attention
 - A modifié la conception du traitement en cancérologie après quelques succès importants
 - A fait naître des espoirs importants pour les patients, les médecins, ...
- A conduit à des investissements considérables
 - Equipes de recherche universitaires
 - Industrie pharmaceutique
 - *Precision Medicine Initiative*

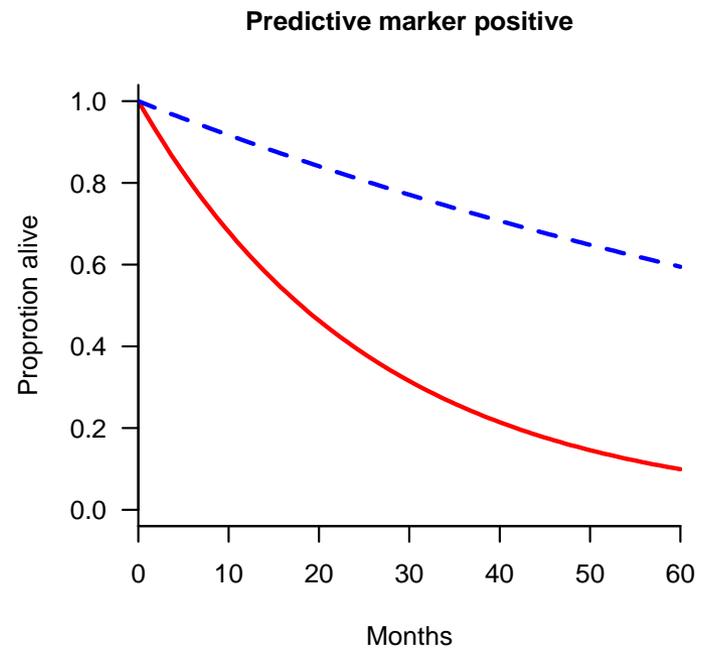
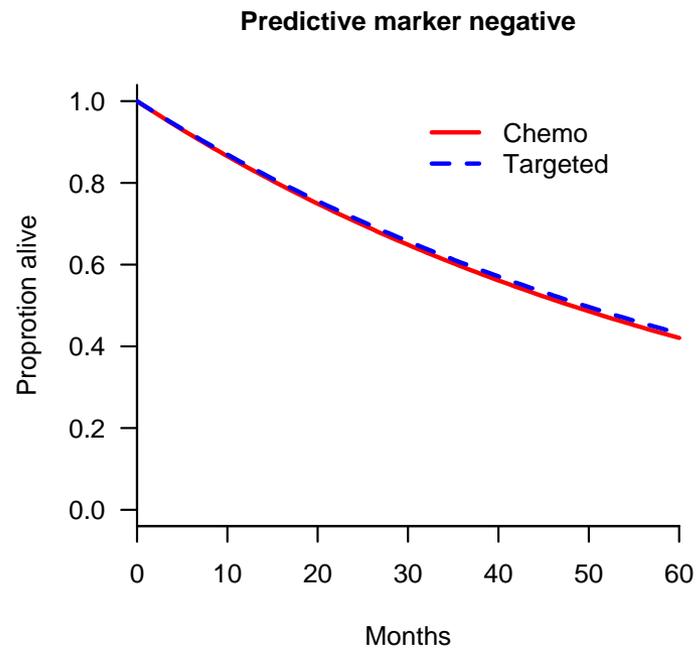
“Tonight, I’m launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes — and to give all of us access to the personalized information we need to keep ourselves and our families healthier.”

— President Barack Obama, State of the Union Address, January 20, 2015

Qu'est-ce?

- Repose sur l'utilisation de biomarqueurs pour guider le choix d'un traitement
- Pour maximiser l'efficacité et diminuer la toxicité du traitement
- Exemple: cetuximab dans le cancer colorectal métastatique **non muté KRAS**
- Biomarqueur: plusieurs définitions
 - Altération moléculaire mesurable qui définit un phénotype biologique caractérisant un état physiologique ou pathologique
- En fait, la notion de biomarqueur peut être étendue
 - Toute combinaison de biomarqueurs mais aussi d'autres tests biologiques, d'imagerie, de caractéristiques d'un patient
 - Qui peuvent être utilisées comme outil diagnostique, pronostique ou **prédictif**

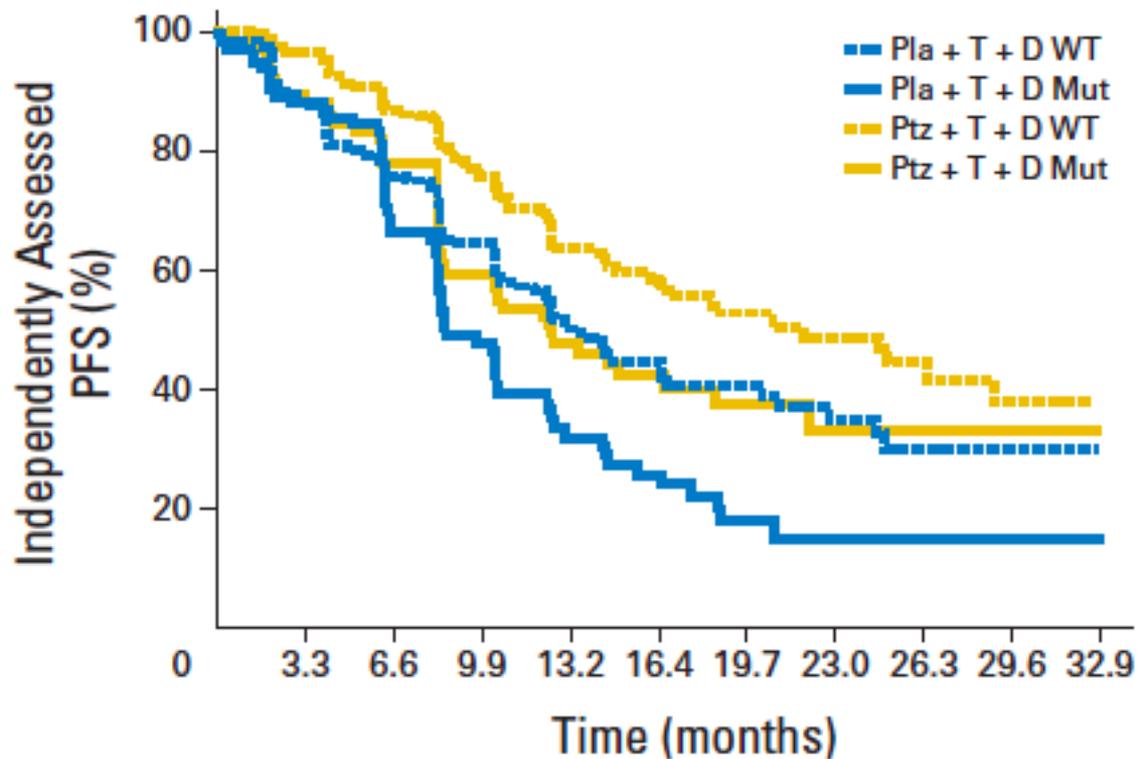
Principe de la médecine de précision



Biomarqueurs de (pour) quoi?

- Diagnostic
- Pronostic
- Suivi de la réponse
- **Prédiction de la réponse**
 - Marqueurs associés à la réponse au traitement
 - En termes d'efficacité ou de toxicité
- Bien faire attention à la différence marqueur pronostique / prédictif de réponse

Exemple d'un biomarqueur pronostique



Cancer du sein métastatique
HER2+

Trastuzumab + docetaxel +
pertuzumab ou placebo

PIK3CA sauvage (WT) ou
muté (Mut)

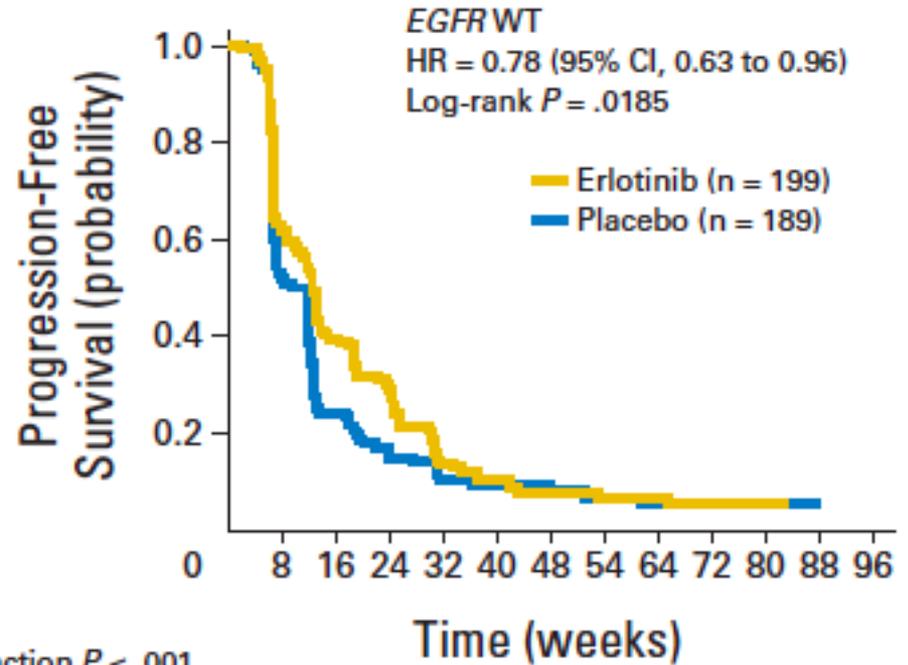
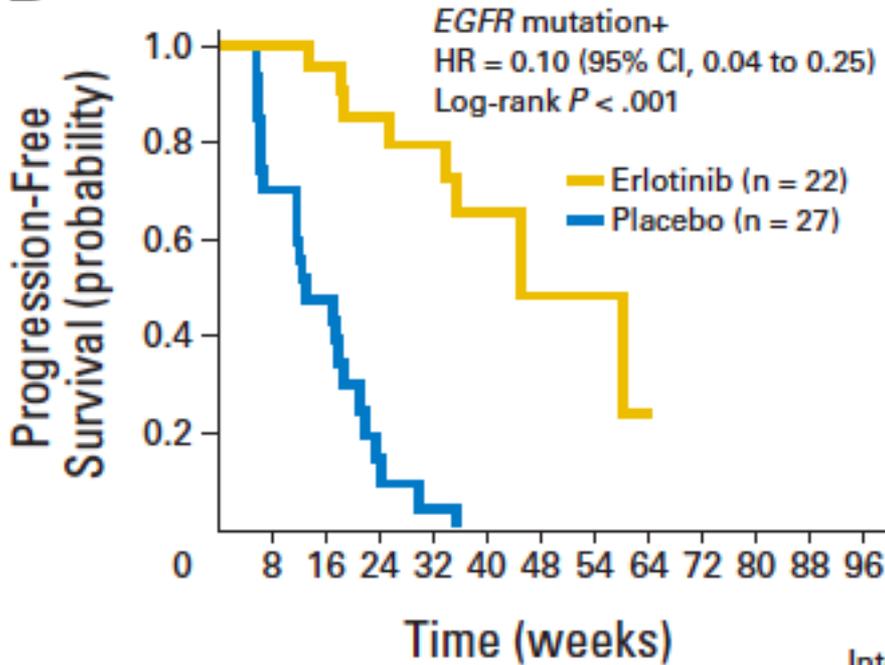
No. at risk											
Pla + T + D WT	191	164	136	114	66	46	23	17	9	3	1
Pla + T + D Mut	90	76	56	37	21	17	8	4	3	2	1
Ptz + T + D WT	190	179	159	137	90	71	46	26	16	5	3
Ptz + T + D Mut	86	71	61	44	29	25	12	6	2	1	1

© 2015 by American Society of Clinical Oncology

Ballman J Clin Oncol 2015; 33:3968-3972

Exemple d'un biomarqueur prédictif

B



Interaction $P < .001$

© 2015 by American Society of Clinical Oncology

Cancer bronchique non à petites cellules
Erlotinib vs placebo
EGFR muté ou sauvage (WT)

Identification des répondeurs à un traitement

- Essais cliniques avec des plans d'expériences particuliers (*enriched, biomarker-stratified, biomarker-strategy* et *adaptive biomarker designs*)
- Méthodes « *data-driven* » pour identifier et confirmer des sous-groupes de répondeurs (cette présentation)
 - Méthodes plus exploratoires
 - On essaie cependant de contrôler le risque d'erreur de première espèce

Pourquoi / quand utiliser cette approche ?

- Un essai ne montre pas de bénéfice dans l'ensemble de la population → peut-être qu'un sous groupe de patients en bénéficient quand-même
- Un essai montre un bénéfice, mais celui-ci est peut-être entièrement dû à un nombre plus restreint de très bons répondeurs
- Un essai montre un bénéfice, mais il est aussi toxique(ou coûte très cher) → restreindre sa prescription aux patient pour qui le bénéfice dépasse le risque (ou le coût)

Analyses en sous-groupes?

- Solution « historique » pour trouver des sous-groupes de répondeurs à un traitement, ou de patients avec une réponse plus importante
 - Peu d'avantages et inconvénients importants
 - Avantage: assez facile à faire et méthode classique
 - Inconvénients: multiplicité généralement non prise en compte, approche « variable par variable », optimisme dans les estimations, ...
- Nécessité d'avoir des méthodes statistiquement valides pour définir des sous-groupes de patients répondeurs

Principes statistiques à considérer

- Contrôle du risque d'erreur de première espèce de l'ensemble de la stratégie
 - Contrôle du risque d'erreur de type I **au sens fort**
 - Ou contrôle du taux de fausses découvertes (**FDR**), selon les objectifs
- Contrôle de la complexité du modèle pour éviter le sur-ajustement
- Evaluation de la reproductibilité des sous-groupes identifiés
- Estimation sans biais de l'effet du traitement dans le(s) sous-groupe(s) identifié(s)
- Prise en compte de l'incertitude sur l'ensemble de la procédure de détection de sous-groupes

Approche contrefactuelle

- Cadre classique pour l'inférence causale
- Utile aussi pour ce qui nous intéresse ici
- Deux traitements, 1 variable $Z=0$ (contrôle) ou 1 (expérimental)
- Critère de jugement Y
- Pour un patient i , on suppose qu'il existe deux valeurs du critère de jugement Y
- $Y_i(0)$ et $Y_i(1)$ selon que le patient reçoit $Z_i=0$ ou $Z_i=1$
- On peut ainsi définir l'effet causal individuel du traitement

$$D_i = Y_i(1) - Y_i(0)$$

Inférence dans le modèle contrefactuel

- Essai contrôlé randomisé: Z tiré au sort
- On peut alors estimer $E[Y(z)]$ par $E[Y | Z=z]$
- Et plus généralement $f(x,z)=E[Y(z) | X=x]$ par $E[Y | Z=z, X=x]$
- Sous l'hypothèse « SUTVA » (stable unit treatment value)
 - Le traitement donné à un patient n'affecte pas le couple $\{Y(0), Y(1)\}$ des autres patients (pas d'interférence)
 - Il n'existe pas de versions « cachées » du traitement (si $Z=z$, on observe bien $Y(z)$, qui s'appelle l'hypothèse de cohérence)
- On définit aussi le contraste $e(X)$ (effet du traitement au niveau individuel)
$$e(X)=f(X,1) - f(X,0)$$
- Soit $f(X,Z)=f(X,0) + e(X)Z$

Classification (taxonomie) des méthodes

- Méthodes qui modélisent chacun des $f(x,z)$ (*global outcome modelling, GOM*)
- Méthodes qui cherchent à modéliser directement le contraste $e(X)$ (*global treatment effect modeling, GTEM*)
- Méthodes qui se concentrent directement sur la recherche de sous-groupes de patients bénéficiant du traitement (valeurs de $e(X)$ élevées), dites méthodes de modélisation locale (*local modeling, LM*)
- Pour les deux premières catégories, de nombreuses méthodes ont été présentées récemment → choix de présenter quelques unes plutôt qu'un panorama exhaustif

Méthodes *global outcome modeling*

- Méthodes proposées à partir de l'estimation séparée de $f(x,1)$ et $f(x,0)$ dans chacun des groupes, ou avec un seul modèle pour $f(x,z)$ comprenant les interactions covariables x tmt
- Utilisent souvent des méthodes d'apprentissage automatique (*machine learning*) couplée à des approches pénalisées (comme le LASSO)
- Focus sur la méthodes des « jumeaux virtuels » (*virtual twins*, VT) de Foster et coll.

VT

- Méthode en deux étapes
- 1^{ère} étape
- Estimation de $f(x,0)$ et $f(x,1)$ en utilisant des forêts aléatoires avec comme covariables : X_i , Z_i , $X_i \times I(Z_i=0)$ et $X_i \times I(Z_i=1)$
- On en déduit $d_i = \hat{f}(x_i,1) - \hat{f}(x_i,0)$ qui estime l'effet traitement individuel
- 2^e étape
- Utilisation d'arbres de régression ou de classification (CART) sur les d_i pour identifier le sous-groupe S de patients

VT: 2e étape

- 2 choix
- Arbre de régression pour d_i , puis sélection des nœuds où les prédictions sont supérieures à un seuil δ prédéfini comme éléments de S
- Définir une variable binaire $u_i = I(d_i > \delta)$, et utiliser un arbre de classification pour les u_i . S est la réunion des nœuds terminaux prédisant $u_i=1$

VT: Estimation de l'effet du traitement dans S

- Mesure de bénéfice $Q(S)$

$$Q(S) = E(Y | Z = 1, X \in S) - E(Y | Z = 0, X \in S) - [E(Y | Z = 1) - E(Y | Z = 0)]$$

- On cherche à estimer $Q(\hat{S})$, utilisant le sous-groupe trouvé
- 5 méthodes proposées
 1. Re-substitution (plug-in) des estimations dans l'échantillon
 2. Simulation de nouvelles données (bootstrap paramétrique)
 3. Validation croisée (CV) pour $\hat{f}(x_i, 1)$ et $\hat{f}(x_i, 0)$
 4. Validation croisée de l'ensemble du processus de sélection
 5. Correction du biais par bootstrap (appliquée à l'estimateur obtenu par 1 ou par 2) → meilleures performances que les autres à partir de simulations

Méthodes *Global treatment effect modeling*

- A nouveau plusieurs méthodes
- Parfois considérées comme plus robustes que les précédentes
- Parce que les modèles précédents « consomment » de l'information pour estimer à la fois l'effet des variables purement pronostiques et l'effet de celles prédictives de l'effet du traitement
- Les méthodes GTEM se concentrent sur ces derniers effets → gain d'efficacité
- Illustration: arbres d'interaction (*interaction trees, IT*)

IT

- Extension de CART
- A chaque nœud parent CART estime $y = a_0 + a_1 s + \varepsilon$ où s est l'indicatrice d'une branche candidate, puis un critère pour couper à s construit sur la réduction de la somme des carrés des erreurs
- IT utilise $y = a_0 + a_1 s + a_2 z + a_3 zs + \varepsilon$: interaction tmt x branche (zs)
- Critère pour couper sur la réduction de la somme des carrés des erreurs pour ce terme d'interaction
- Comme dans CART, critères mesurant le coût de la complexité pour déterminer l'arbre final
- IT: coût de la complexité des interactions

Méthodes *Local modeling*

- Moins de méthodes différentes
- Focus sur SIDES
- Trouver un ou plusieurs sous-groupes de patients où l'effet du traitement est sensiblement plus important que dans les données complètes
- Ressemble aussi à CART
- Algorithme récursif dépendant de paramètres:
 - L : nombre maximum de covariables définissant un sous-groupe
 - S : taille minimale d'un sous-groupe
 - M : nombre maximum de covariables candidates à considérer à chaque itération pour définir le sous-groupe enfant

SIDES: étapes de l'algorithme

- A l'itération $k+1$
- Etape 1:
 - Considérer un sous-groupe identifié par la procédure à l'itération k (ensemble des patients si $k=0$)
 - Il sera désigné comme le **groupe parent**
- Etape 2:
 - Pour chaque covariable non encore utilisée, considérer chaque division possible du groupe parent en 2 **sous-groupes enfants**
 - Evaluation du critère de « séparation » afin de garder les M meilleures paires de sous-groupes enfants
 - Pour chacune, garder le sous-groupe avec le meilleur effet traitement comme **sous-groupe prometteur**
 - Si la taille du sous-groupe est inférieure à S , on l'exclut
- Etape 3:
 - Evaluation des critères de « sélection » et de « continuation » pour classer les sous-groupes prometteurs en **sous-groupes candidats** et **sous-groupes parents** pour la prochaine itération

SIDES: Critère de séparation

- Outil pour la sélection de sous-groupes enfants avec un meilleur effet traitement
- On considère 2 sous-groupes enfants. Soit Z_{E1} et Z_{E2} les statistiques de test de l'hypothèse de l'efficacité du traitement
 - Maximisation de l'effet traitement différentiel
$$p_1 = 2 \left[1 - \phi \left(\frac{|Z_{E1} - Z_{E2}|}{\sqrt{2}} \right) \right]$$
 - Maximisation de l'effet traitement dans au moins un des 2 sous-groupes
$$p_2 = 2 \min \left(1 - f(Z_{E1}), 1 - f(Z_{E2}) \right)$$

Où ϕ est la fonction de répartition d'une loi normale standard

SIDES: Critère de continuation

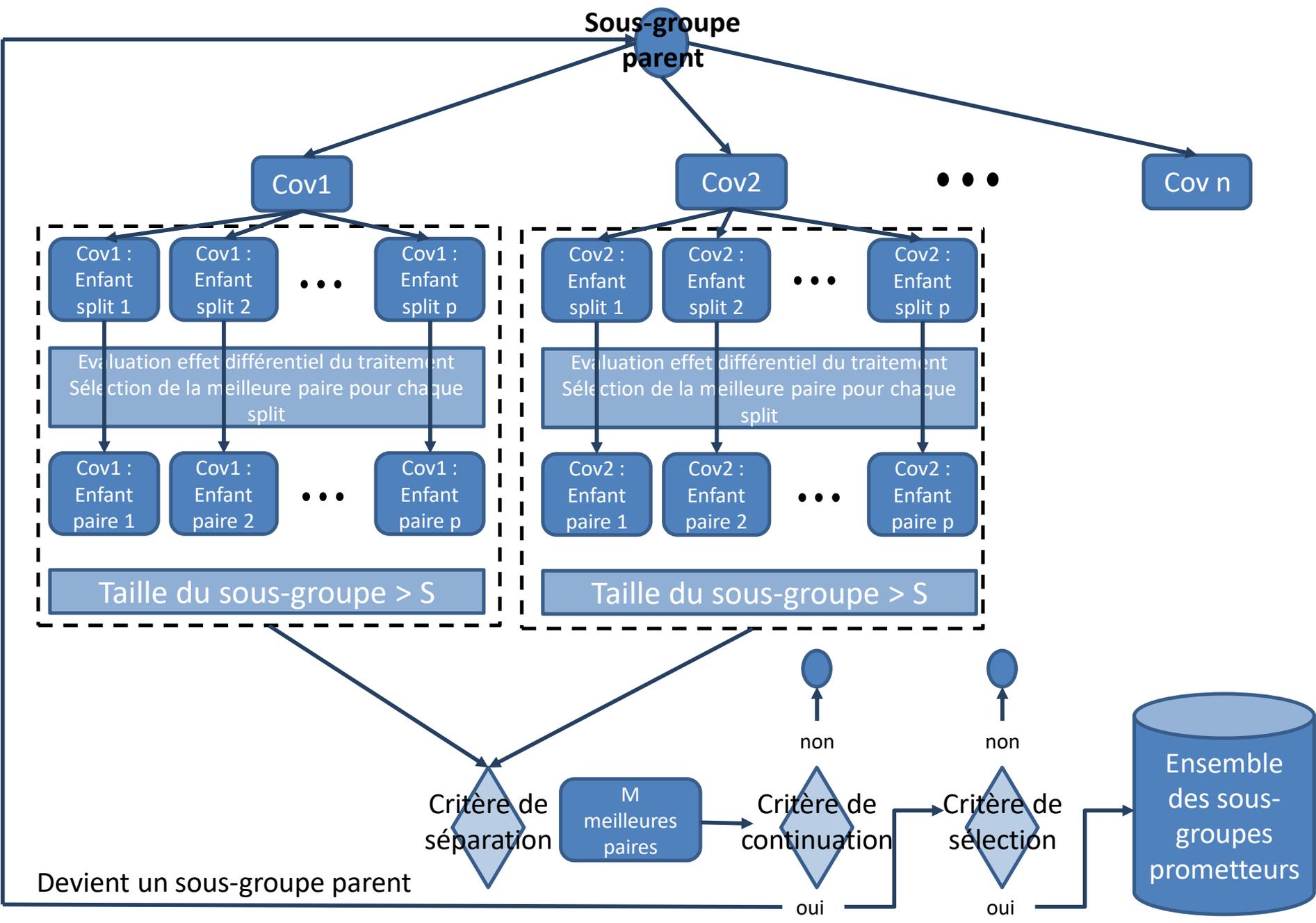
- Aide à contrôler le nombre de sous-groupes à évaluer par la procédure d'identification
- Conservation des sous-groupes qui ont une amélioration comparé à leur parent
- Se fonde sur les p-valeurs de l'effet traitement du sous-groupe enfant (p_c) et du groupe parent (p_p)
- Un sous-groupe prometteur devient un parent seulement si

$$p_c \leq \gamma p_p$$

- Où $0 < \gamma \leq 1$ est le paramètre d'amélioration relative, qui peut être déterminé de façon optimale par validation croisée

SIDES: Critère de sélection

- Défini la règle pour identifier un sous-groupe candidat
- Contrôle l'erreur globale de type I au sens faible
- Un sous-groupe devient candidat lorsque sa p -valeur est significative au niveau α , trouvé grâce à une méthode de ré-échantillonnage:
 - Permutation entre les covariables et le vecteur des Y_i
 - Faire un grand nombre de permutations
 - Définir des valeurs de coupure α_j
 - Pour chaque permutation, compter le nombre de fois où le critère de sélection est vérifié pour au moins un sous-groupe
 - Sélectionner la plus grande valeur $\alpha_j \leq 5\%$



Sous-groupe parent

Cov1

Cov2

...

Cov n

Cov1 :
Enfant
split 1

Cov1 :
Enfant
split 2

...

Cov1 :
Enfant
split p

Cov2 :
Enfant
split 1

Cov2 :
Enfant
split 2

...

Cov2 :
Enfant
split p

Evaluation effet différentiel du traitement
Sélection de la meilleure paire pour chaque
split

Evaluation effet différentiel du traitement
Sélection de la meilleure paire pour chaque
split

Cov1 :
Enfant
paire 1

Cov1 :
Enfant
paire 2

...

Cov1 :
Enfant
paire p

Cov2 :
Enfant
paire 1

Cov2 :
Enfant
paire 2

...

Cov2 :
Enfant
paire p

Taille du sous-groupe > S

Taille du sous-groupe > S

Critère de séparation

M
meilleures
paires

Critère de continuation

Critère de sélection

Ensemble
des sous-
groupes
prometteurs

Devient un sous-groupe parent

non

non

oui

oui

Illustration

- Essai de phase III dans le sepsis sévère
- 2 bras parallèles, randomisation 2:1 entre tmt expérimental vs soins habituels (N=317 / 153)
- Critère de jugement principal: survie à 28 jours
- Résultats: survie moins bonne dans le bras expérimental (p=0.83, test unilatéral)
- 11 biomarqueurs candidats (tous variables continues)
 - Clinique: Age, délai d'administration du médicament, activité à baseline, nombre de défaillances d'organes
 - Scores: SOFA, APACHE II, GLASGOW
 - Biologie: plaquettes, créatinine, IL-6, bilirubine

Illustration: VT

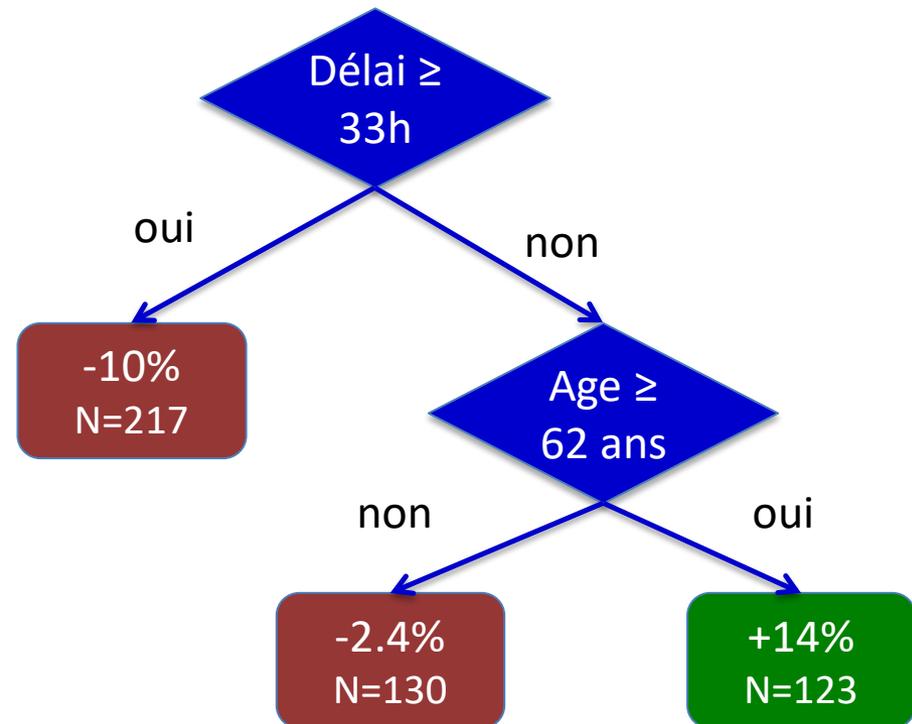
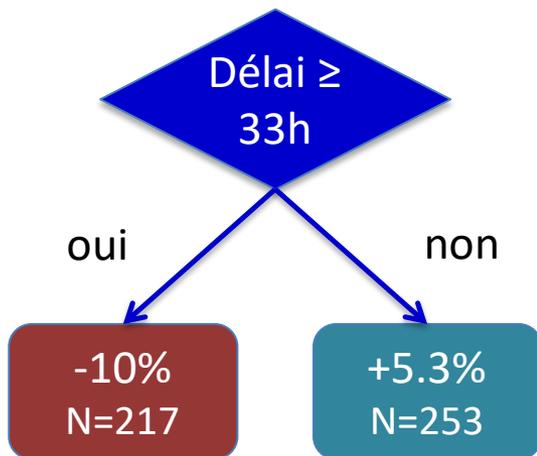
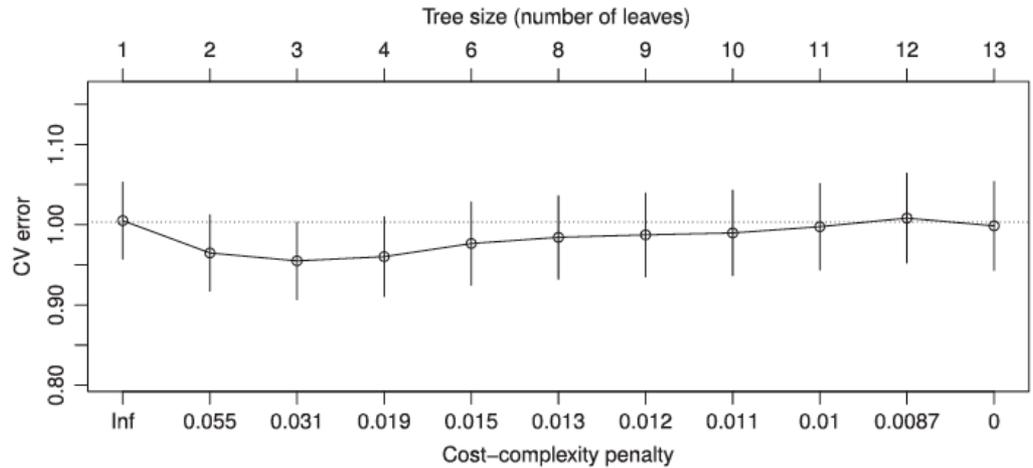


Illustration: SIDES

- Extension SIDESscreen: 2 étapes:
 - 1A: SIDES sans contrôle de complexité pour produire de nombreux sous-groupes « prometteurs »
 - 1B: Calcul d'un score d'importance de chaque variable (VI)
 - 1C: Seuil de VI pour retenir les biomarqueurs à partir de la distribution nulle des VI
 - 2: SIDES sur le sous-ensemble de marqueurs, avec contrôle de complexité et de multiplicité

		Test unilatéral de l'effet du tmt	
Sous-groupe	Taille	p non ajusté	p ajusté
Délai ≤ 31 h et âge > 60 ans	123	0.002	0.044

Conclusion

- Introduction à ces méthodes sans rentrer dans les détails
- Aller plus loin dans l'explication des méthodes prendrait (beaucoup) plus de temps
- Importance du contrôle de la multiplicité et de la validation des sous-groupes
- Champ récent et très actif de la recherche en biostatistique
- Beaucoup de méthodes (très) récentes et peu d'utilisation
- Nous n'avons pas abordé les stratégies optimales de traitement (*optimal treatment regimes*, OTR)
 - Conceptuellement proche
 - Déterminer **quel traitement est adapté à un patient donné** plutôt que quels patients répondent à un traitement donné

Points à retenir

- Il existe des méthodes de recherche de sous-groupes de patients répondeurs qui respectent des prérequis statistiques
- Méthodes souvent en plusieurs étapes
- Peuvent utiliser des approches statistiques récentes et complexes (*machine learning* ...)
- Etapes de validation, correction pour la multiplicité ou de correction des biais dans l'estimation de l'effet du traitement importantes
- Ces étapes utilisent des approches variées: validation croisée, bootstrap, tests de permutation