

ESSAIS CLINIQUES

PASS Lyon Est

Dr. Nicolas ROMAIN-SCELLE

LA PREUVE D'EFFICACITÉ

Efficacité d'un traitement

- Selon les principes de l'Evidence Based Medecine et de l'éthique médicale, l'efficacité des traitements proposés doit être scientifiquement évaluée
 - On désire prouver que le traitement **impacte** la maladie
- ⇒ Relation **causale** du traitement sur la maladie

Limite de la biostatistique

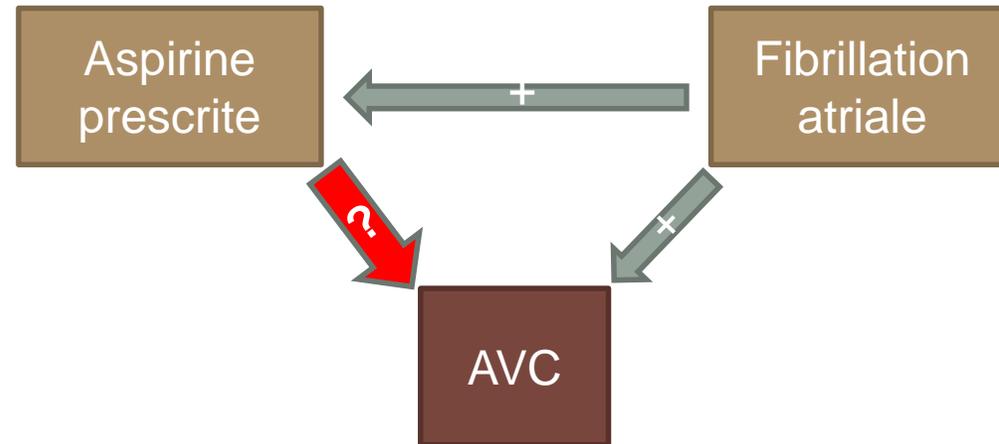
- Un test d'hypothèse isolé met en évidence, au mieux, **une association**
 - **L'approche statistique/mathématique seule ne permet jamais de prouver une causalité**
- ⇒ **Causalité** : nécessité d'une méthode expérimentale particulière

Exemple en soins courant

Quel est l'effet de l'aspirine en prévention de l'AVC chez les patients avec fibrillation atriale ?

Raisonnement :

- AVC : artère bouchée par un thrombus
 - FA : favorise la survenue de thrombus
 - Aspirine : fluidifie le sang
- ⇒ Fluidifier le sang réduit le risque de thrombus donc d'AVC

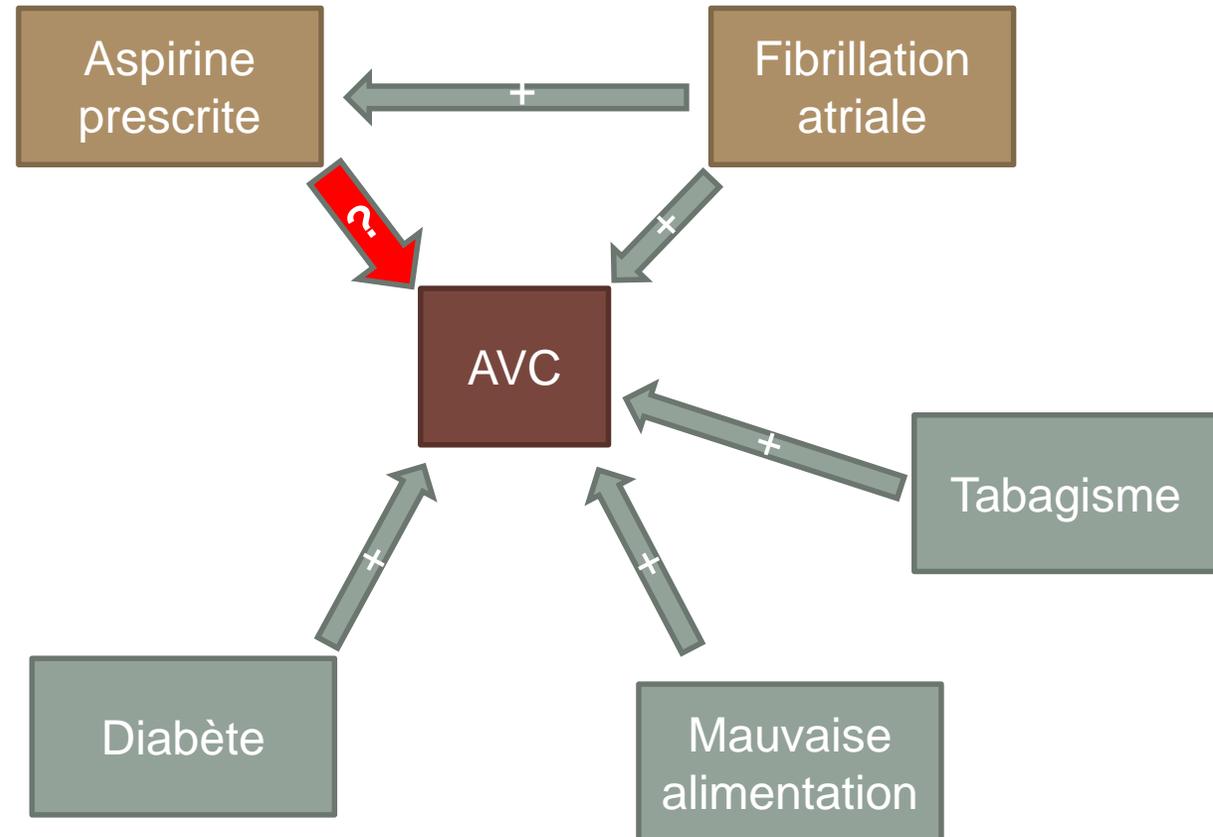


Exemple en soins courant

Quel est l'effet de l'aspirine en prévention de l'AVC chez les patients avec fibrillation atriale ?

Facteurs de risque d'AVC

- Diabète
- Régime alimentaire
- Tabagisme
- **Fibrillation atriale**

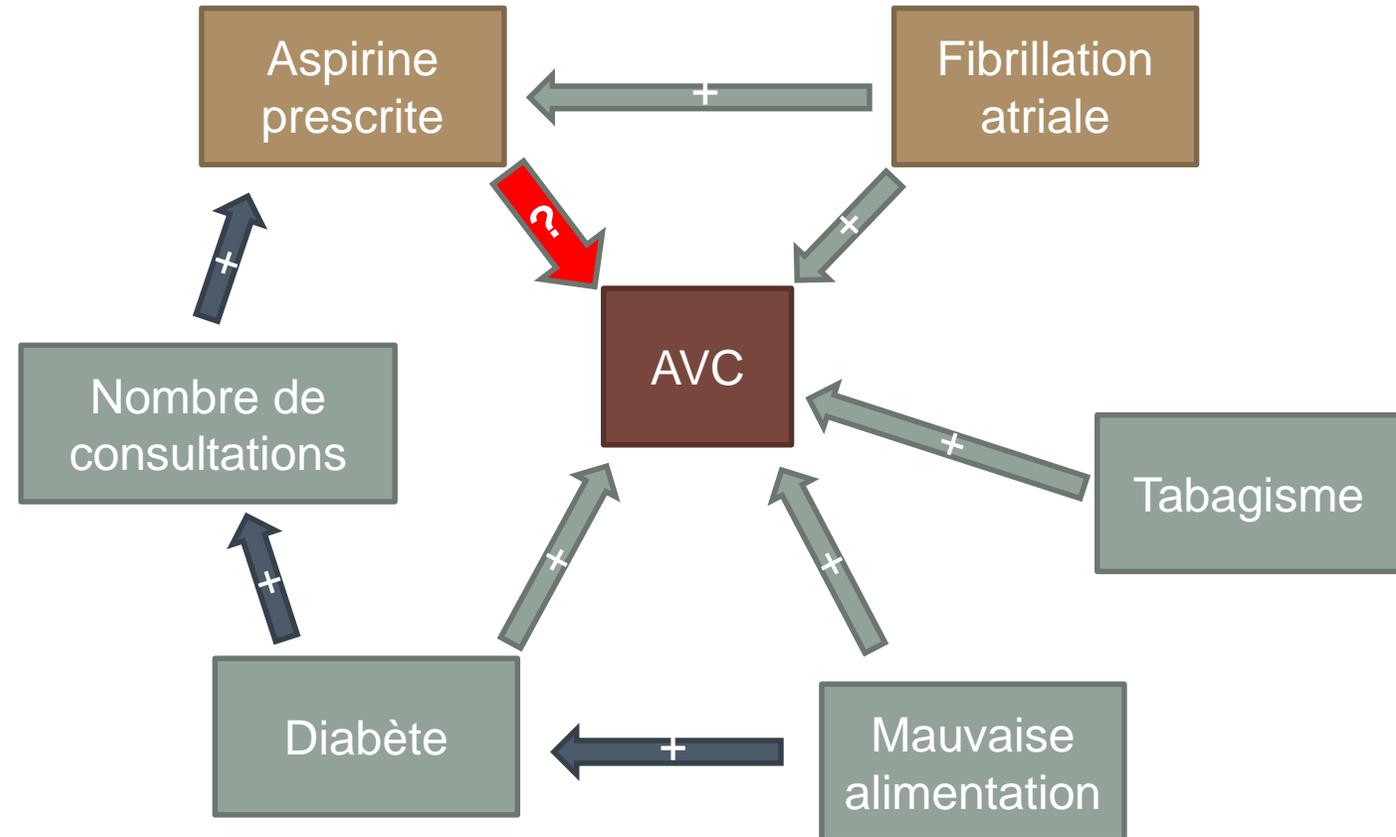


Exemple en soins courant

Quel est l'effet de l'aspirine en prévention de l'AVC chez les patients avec fibrillation atriale ?

Diabète :

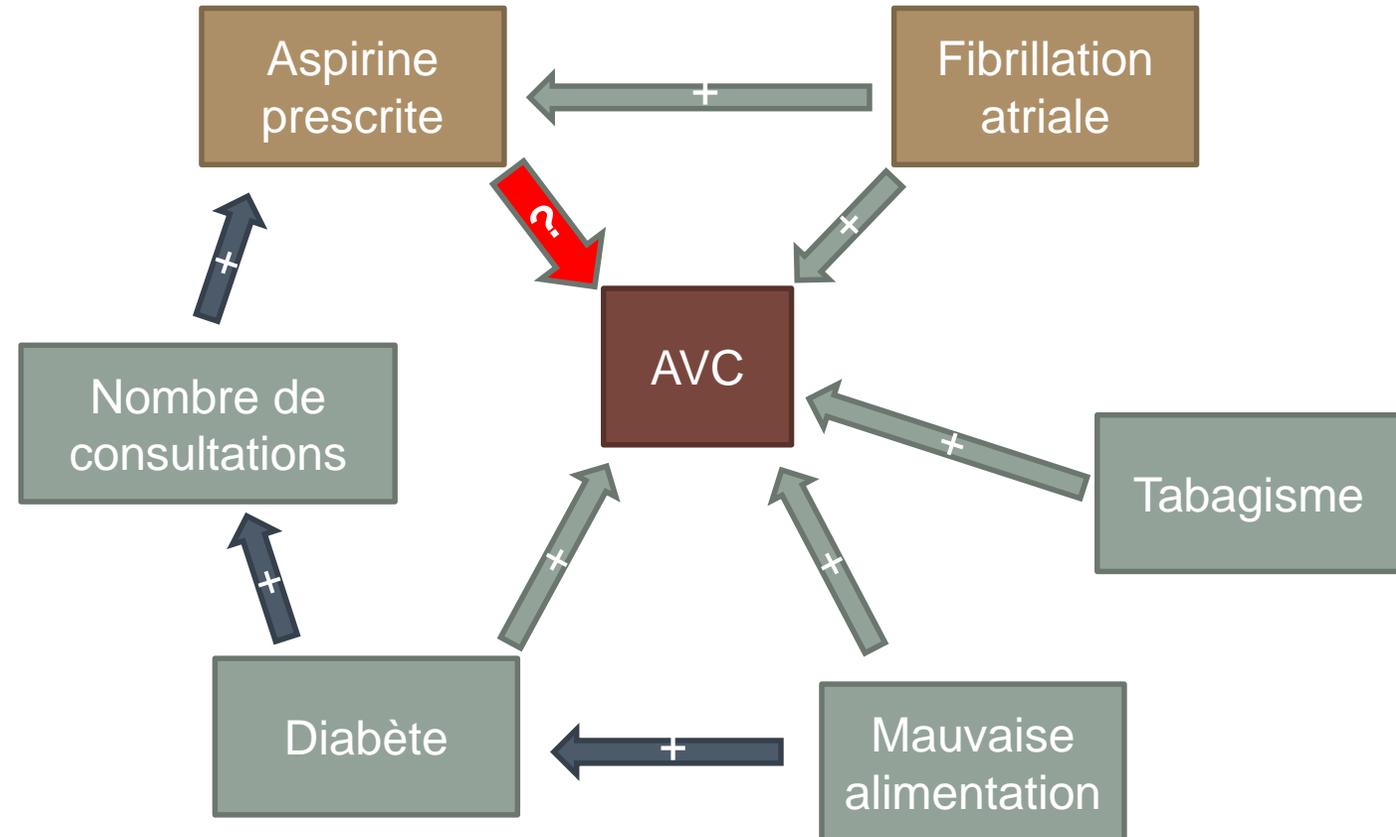
- Dépendant du régime alimentaire
- Suivi médical régulier
- Plus d'opportunités de prescrire l'aspirine



Exemple en soins courants

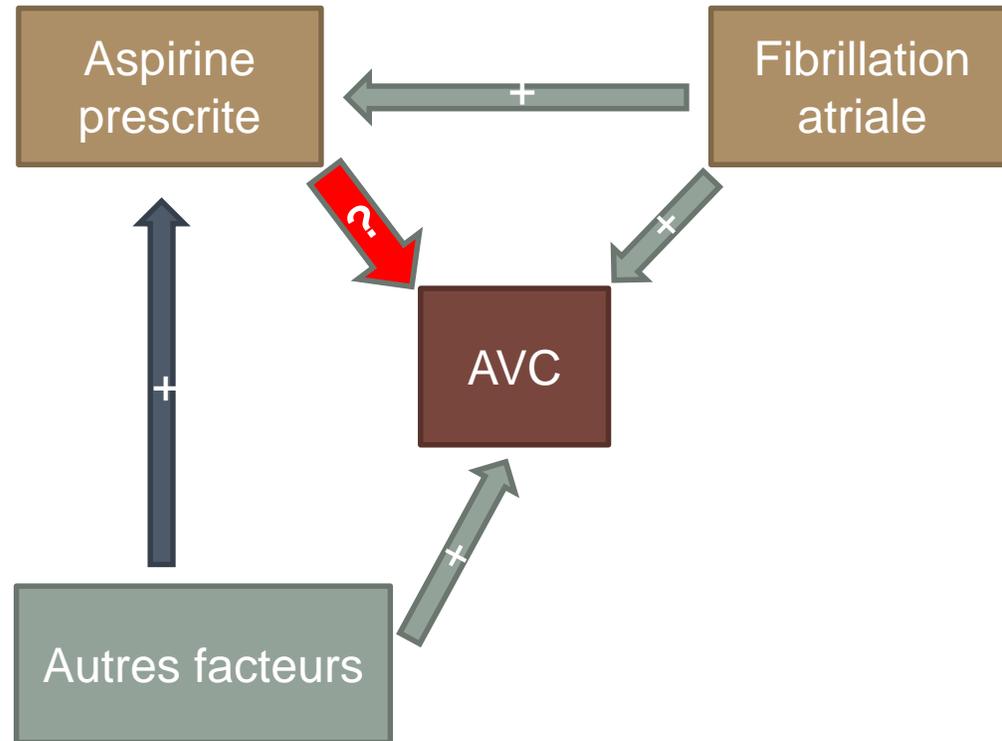
Aspirine : 1 facteur parmi les 5 influençant le risque d'AVC

Comment distinguer le rôle spécifique de l'aspirine parmi les autres facteurs ?



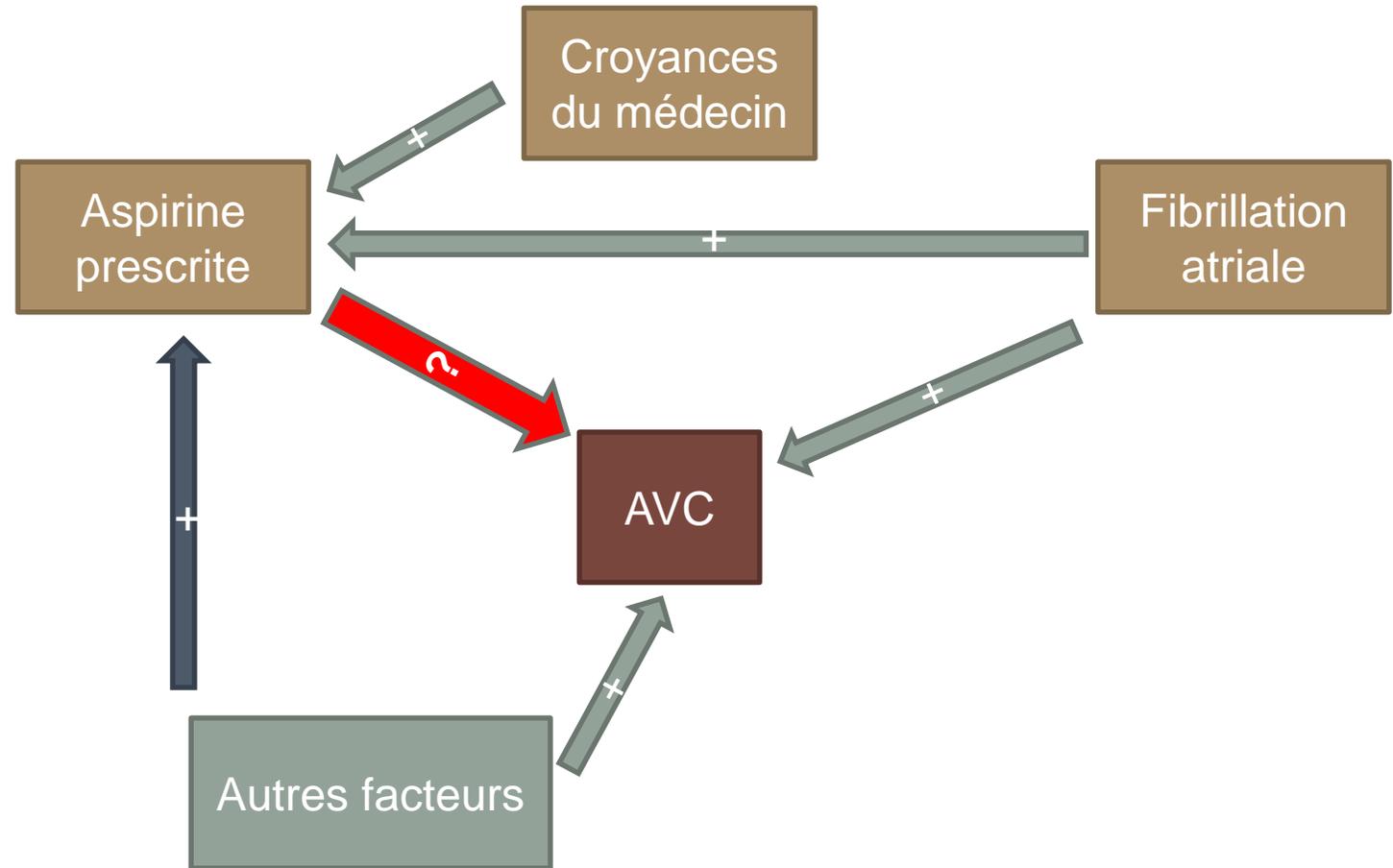
Exemple en soins courant

Simplification : diabète, alimentation et tabac sont placés en autres facteurs, influençant la prescription d'aspirine et le risque d'AVC



Exemple en soins courant

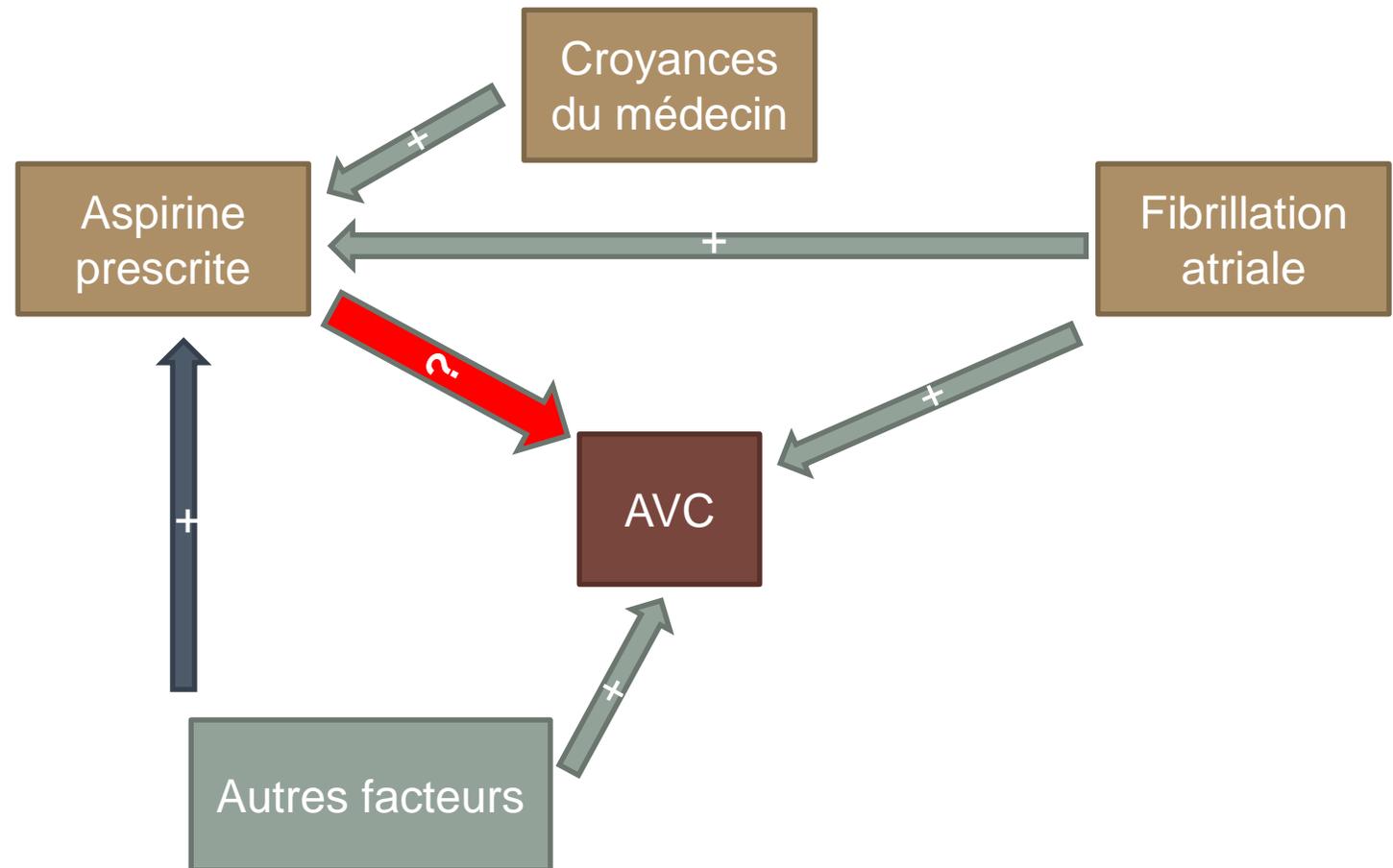
En pratique, en l'absence de preuve, le médecin prescrit l'aspirine selon sa conviction de l'efficacité de l'aspirine



Exemple en soins courants

Si on désire l'effet propre de l'aspirine il faut :

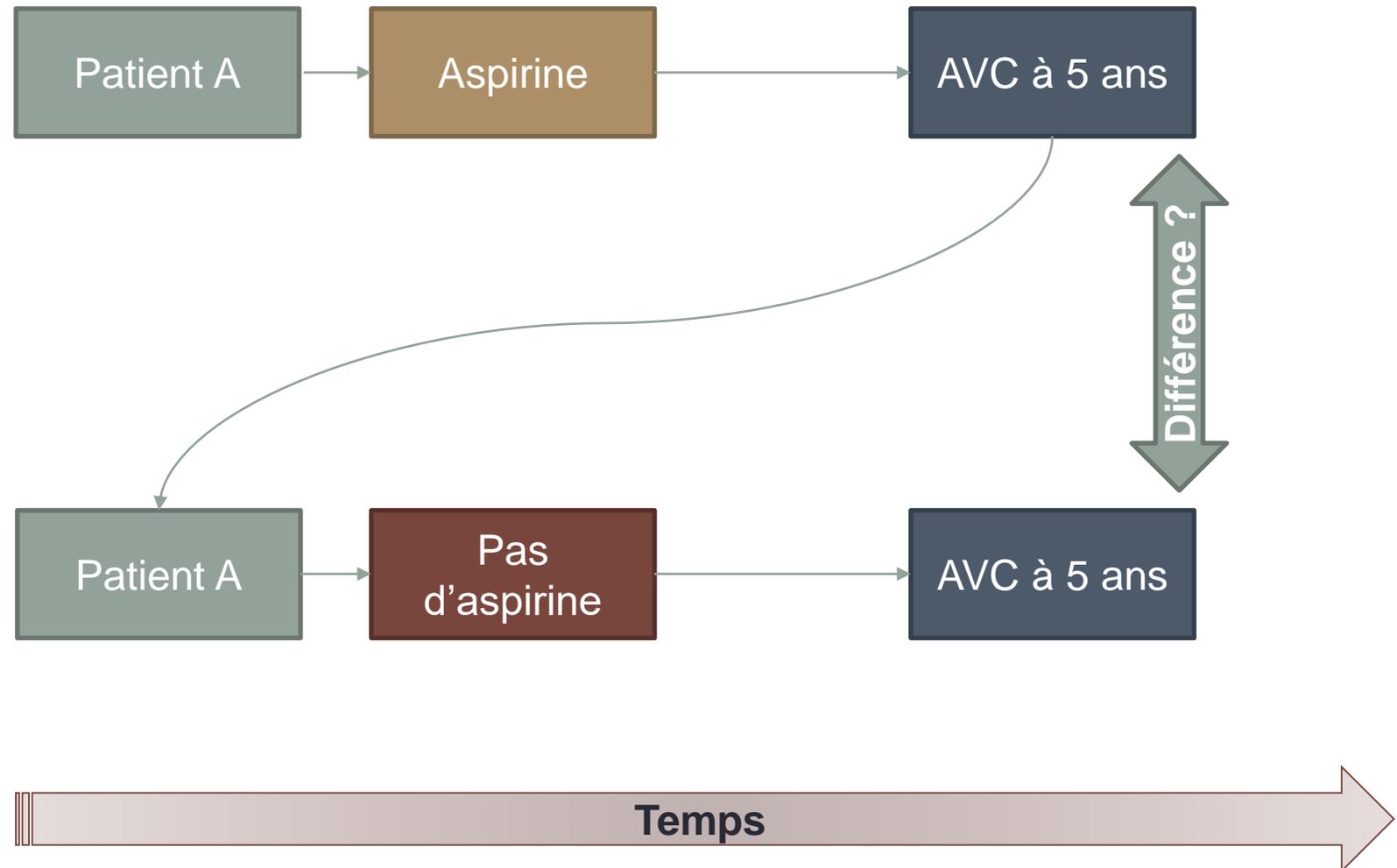
- Supprimer l'effet des croyances du médecin
- Supprimer l'effet de la fibrillation atriale
- Supprimer l'effet des autres facteurs de risque



EXPÉRIENCE IDÉALE

Déterminer l'effet chez un seul patient

- On compare le patient à lui-même, **avec et sans aspirine**
- On regarde, dans chaque cas, la survenue de l'AVC
- On mesure la différence de survenue d'AVC à 5 ans entre les deux suivis

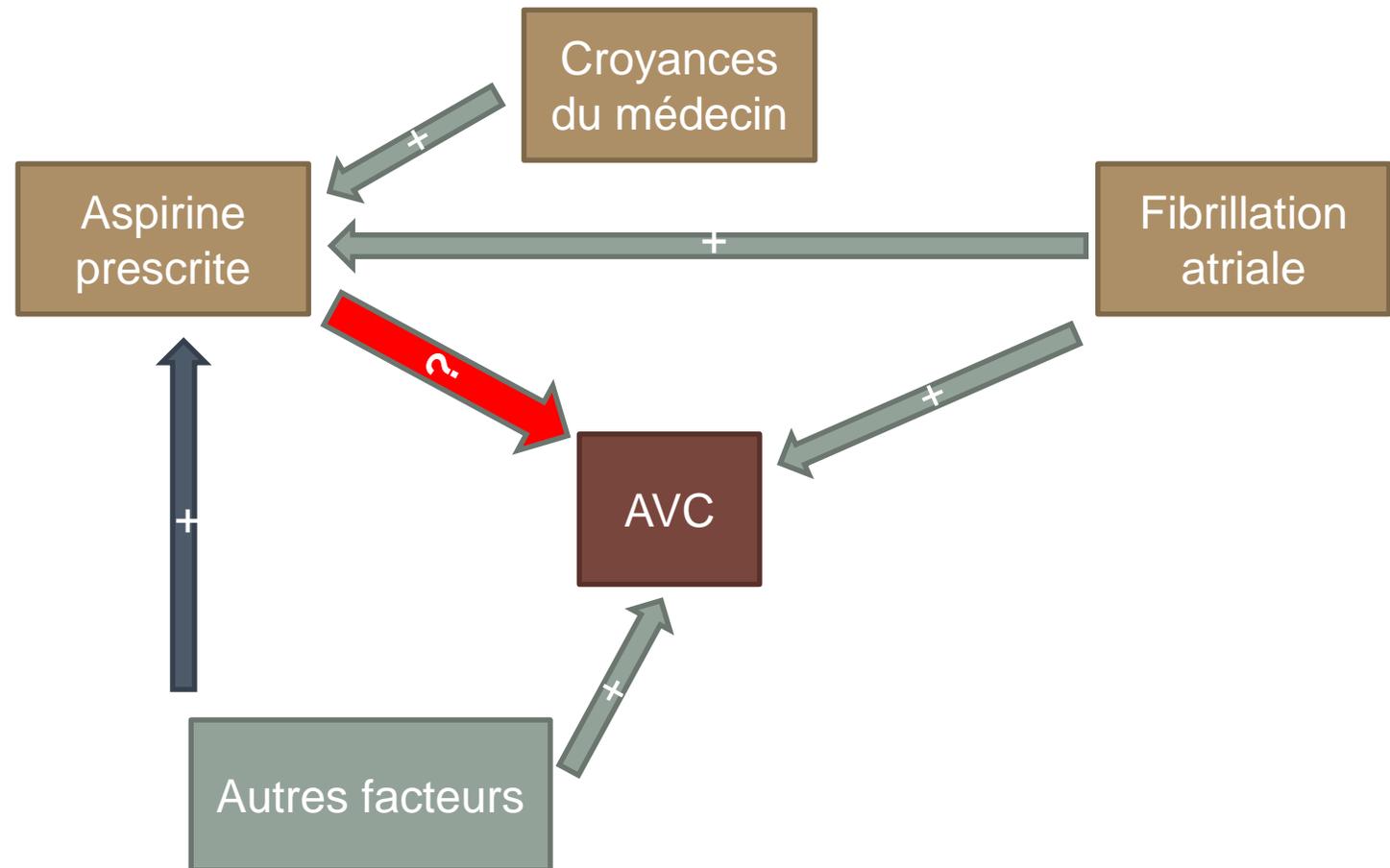


Déterminer l'effet chez un seul patient

En comparant le patient à
lui-même, on supprime
l'effet des autres facteurs

La différence à la fin de
l'expérience est
uniquement due à
l'aspirine

- Avec aspirine

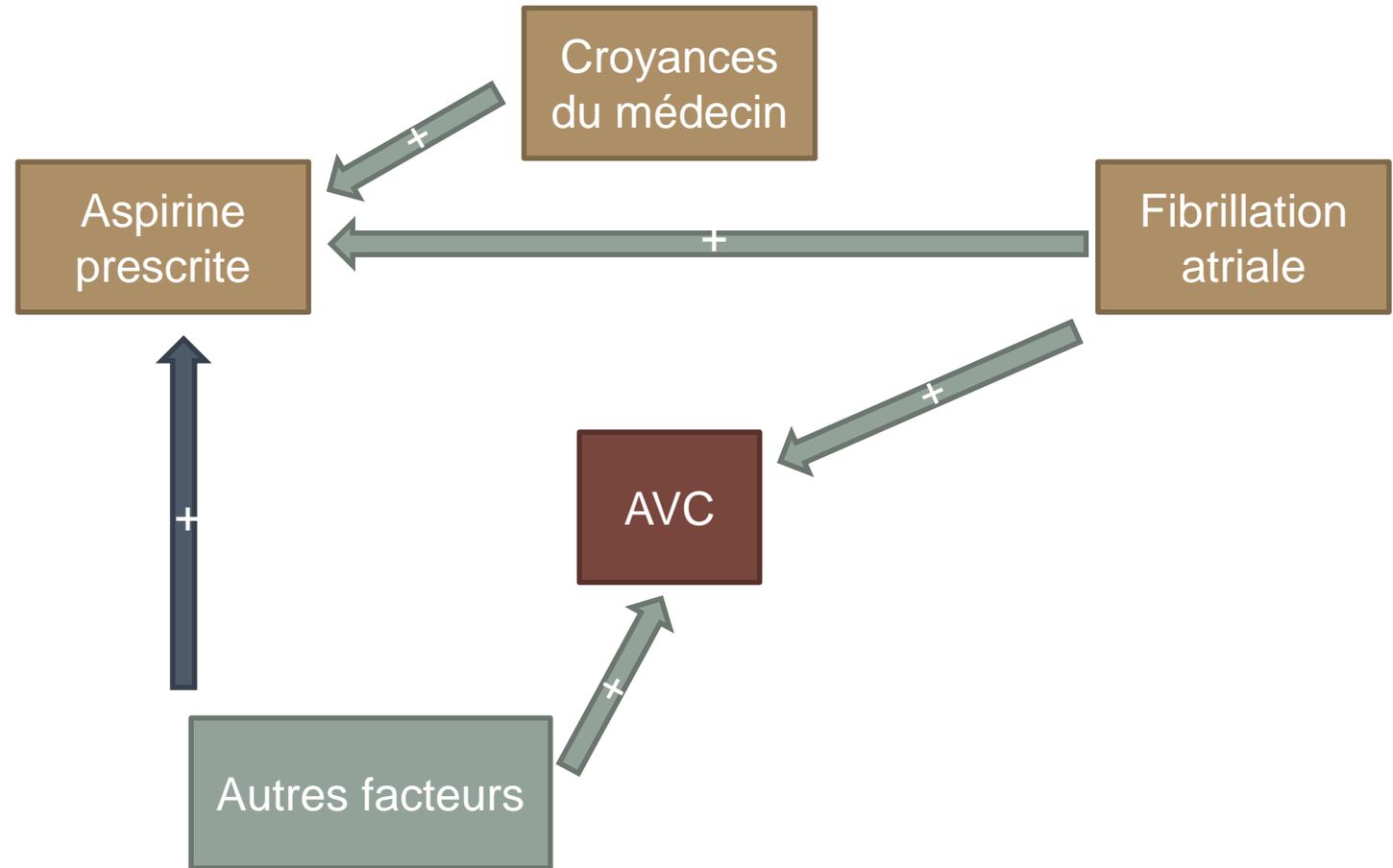


Déterminer l'effet chez un seul patient

En comparant le patient à
lui-même, on supprime
l'effet des autres facteurs

La différence à la fin de
l'expérience est
uniquement due à
l'aspirine

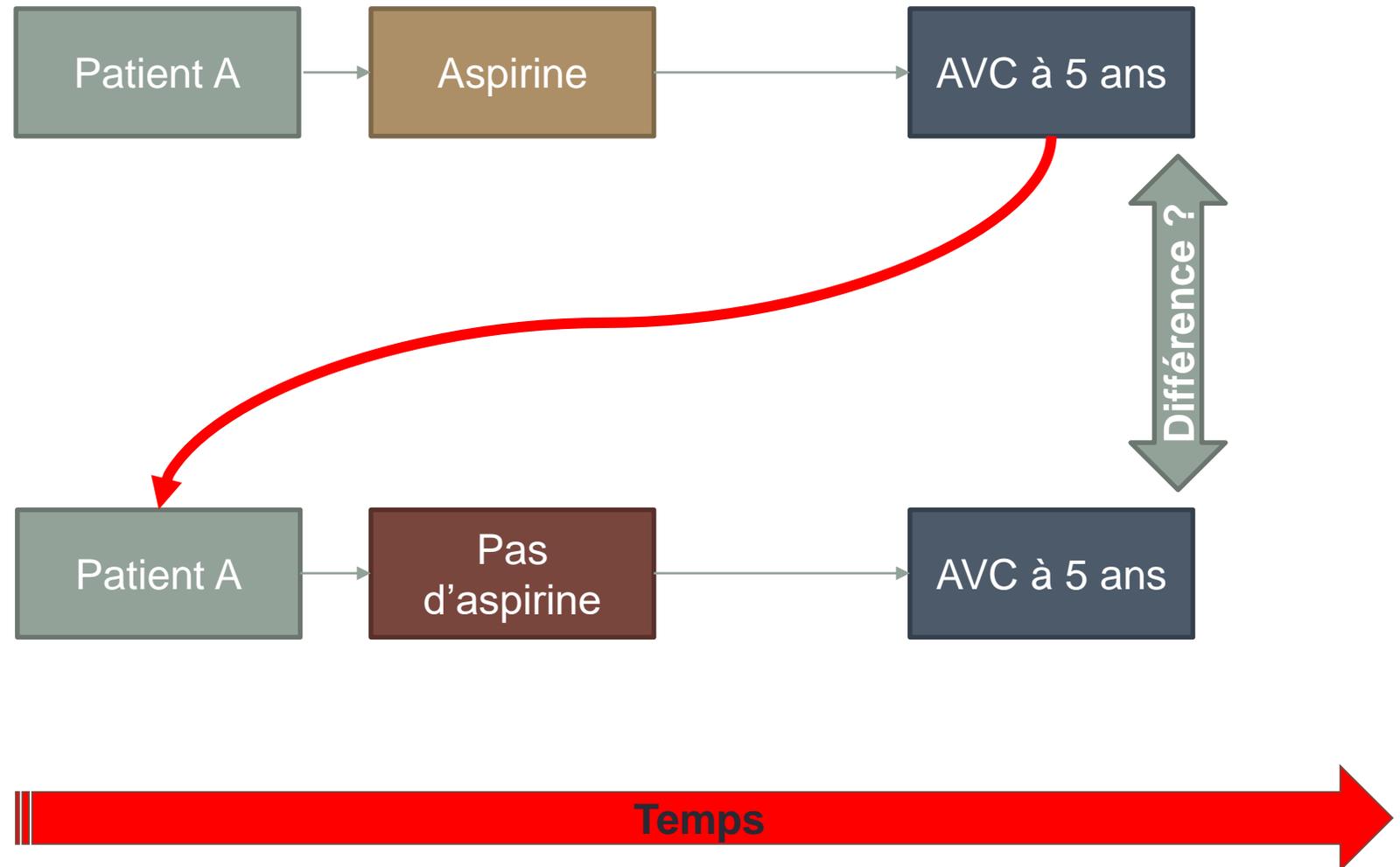
- Sans aspirine



Déterminer l'effet chez un seul patient

Difficulté : **nécessité de remonter le temps**

- ⇒ La comparaison du patient à lui-même est impossible
- ⇒ Nécessité de comparer plusieurs patients entre eux



DÉFINITION DE LA QUESTION

Choisir la question étudiée

- **Population** : quelle est la population de patients dont on souhaite améliorer le traitement (maladie, démographie) ?
- **Intervention** : quel est le nouveau traitement que l'on souhaite évaluer ?
- **Contrôle** : quel est le comparateur adapté au nouveau traitement évalué ?
- **Outcome** (critère de jugement) : sur quel critère va-t-on mesurer l'effet du nouveau traitement comparativement au contrôle ?

Supériorité vs. autres questions

Essai de supériorité

- **Prouver l'efficacité supérieure du nouveau traitement par rapport à la référence en vigueur**
- Majorité des essais : **améliorer les traitements proposés au patient**

Essai de non infériorité/équivalence

- Prouver qu'un nouveau traitement ne fait pas pire ou est équivalent à la référence en vigueur
- Minorité des essais : améliorer la sécurité des traitements proposés au patient

Les autres types d'essais ne sont pas le sujet de ce cours, seule la supériorité est étudiée ici.

SÉLECTION DES PATIENTS

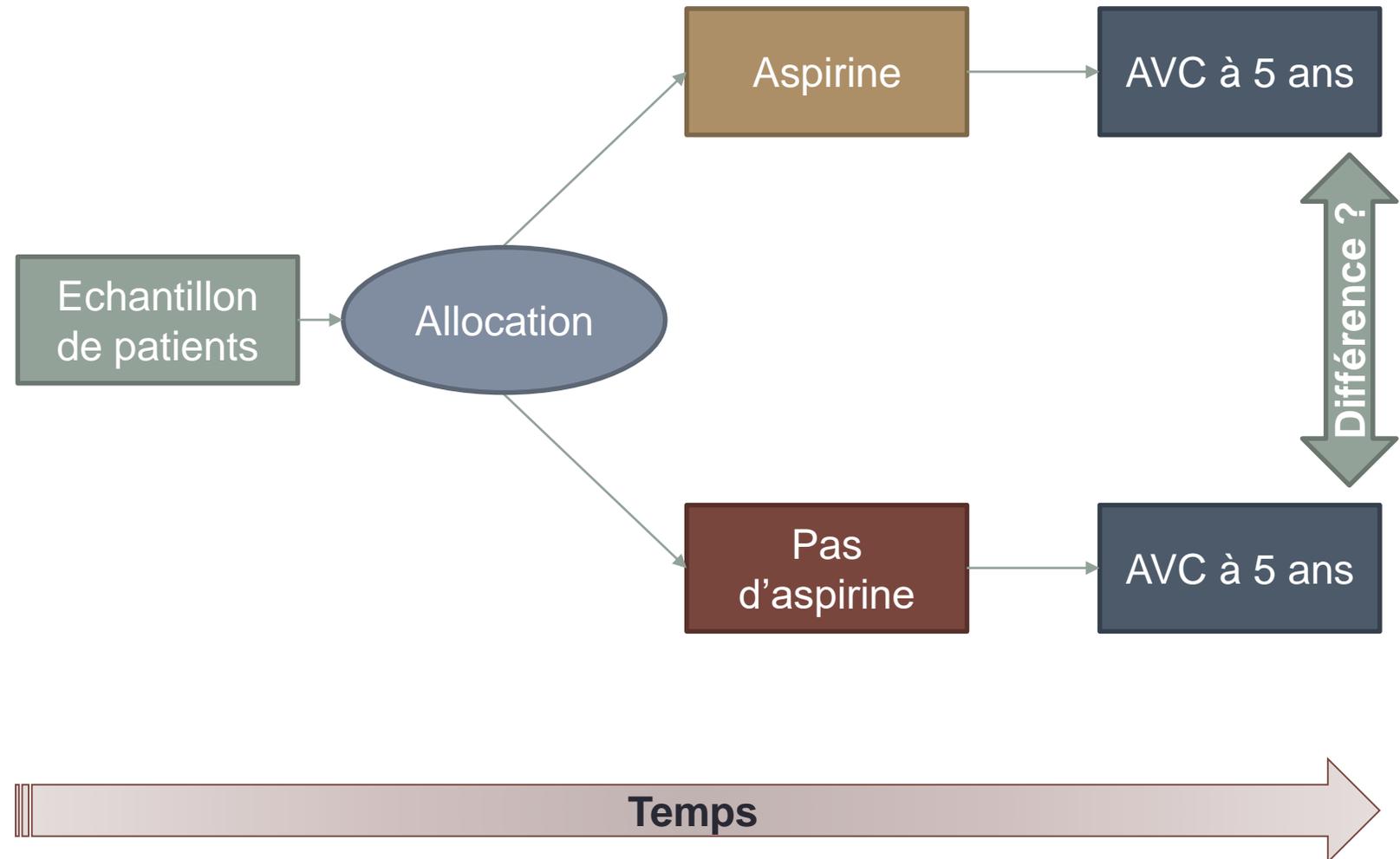
Design expérimental

En pratique, on compare deux groupes de patients entre eux :

- Avec aspirine : **intervention**
- Sans aspirine : **contrôle**

Chaque groupe est appelé un **bras**

On mesure après 5 ans de suivi la différence de fréquence d'AVC



Etape 1 : constituer l'échantillon

- Rappel : effet de l'aspirine en prévention de l'AVC **chez les patients atteints de fibrillation atriale**
- Critères de constitution de l'échantillon :
 - Présence d'une FA
 - Pas de contreindication à l'aspirine
- **Critères d'inclusion et d'exclusion**

Populations en essai clinique

Population cible :

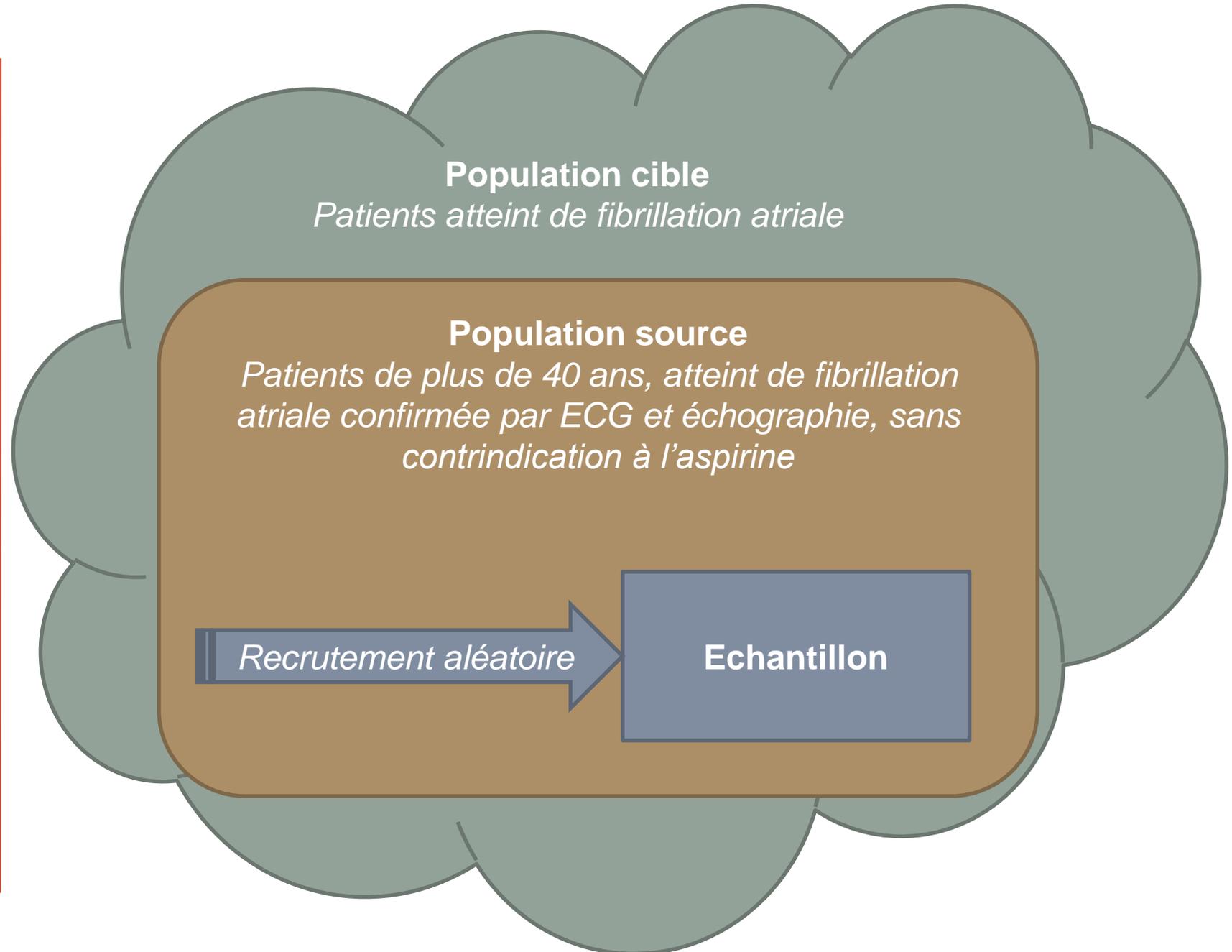
- Population d'intérêt clinique
- Proposer une amélioration

Population source :

- Population définie par l'étude
- Critères d'inclusion et d'exclusion

Echantillon

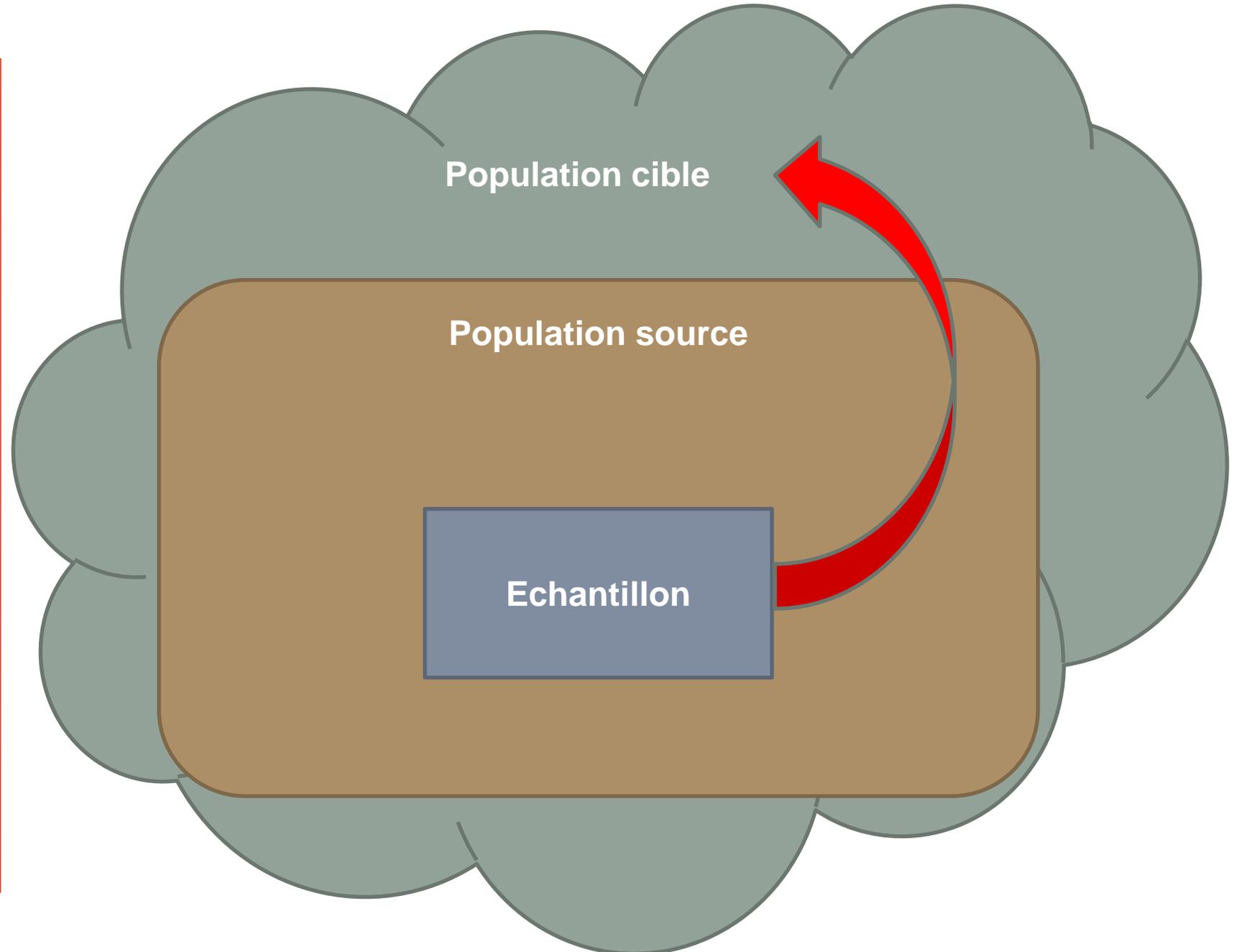
- Recrutement aléatoire dans la population source
- Représentatif



Inférence

On fait une **inférence** sur l'effet du traitement dans la population cible à partir du résultat dans l'échantillon

Les **critères d'inclusion et d'exclusion conditionnent la représentativité** et donc la validité de l'inférence



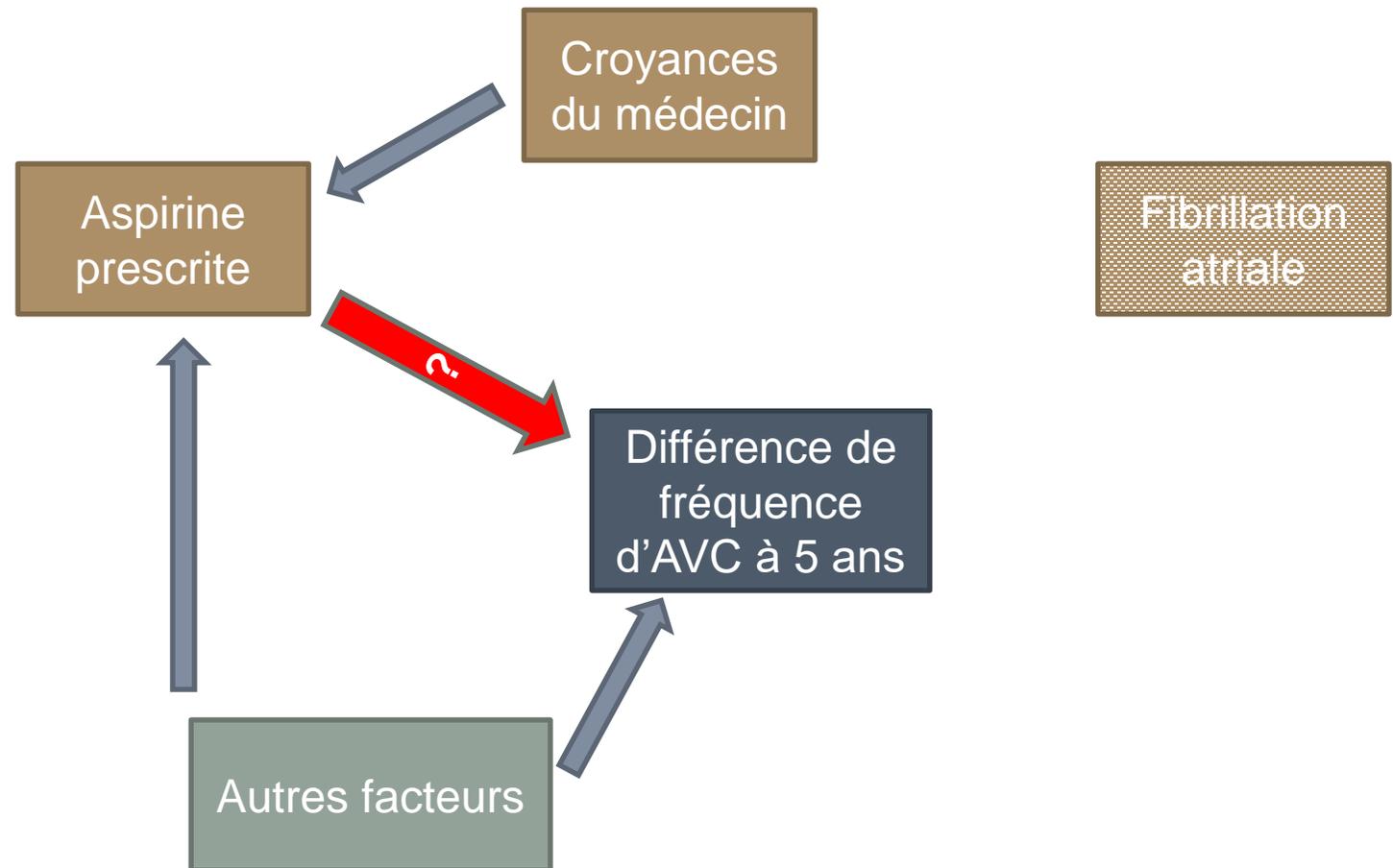
Effet des critères d'inclusion et d'exclusion

Tous les patients étudiés présentent une FA

⇒ Pas de différence de présence de FA dans l'échantillon

⇒ **La FA n'a plus d'effet sur la différence de fréquence d'AVC entre les bras**

• Facteurs influençant la différence



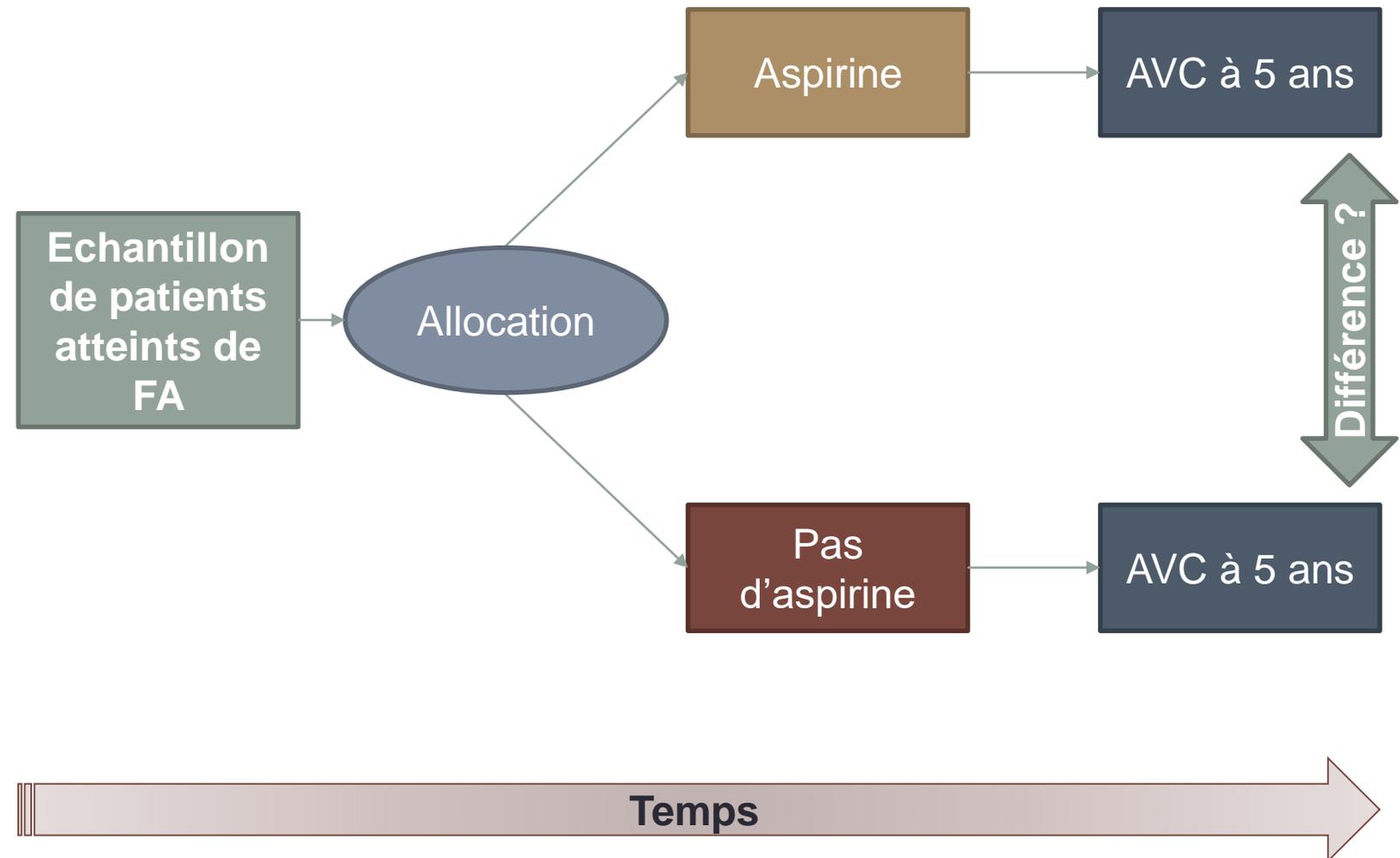
Design expérimental

En pratique, on compare deux groupes de patients entre eux :

- Avec aspirine : **intervention**
- Sans aspirine : **contrôle**

Chaque groupe est appelé un **bras**

On mesure après 5 ans de suivi la différence de fréquence d'AVC



RANDOMISATION

Etape 2 : allouer les patients dans un bras

- Prescription d'un médicament : **choix du médecin**
- Choix partiellement rationnel :
 - Croyances sur l'efficacité du traitement
 - Facteurs inconscients (expérience+++)
- *Exemple : un patient diabétique aura plus de chances de recevoir de l'aspirine qu'un patient non diabétique par un même médecin, celui-ci voulant contrôler plus activement le risque d'AVC déjà élevé*

Etape 2 : allouer les patients dans un bras

- Eliminer les choix implicites et explicites du médecin

⇒ **Randomisation**

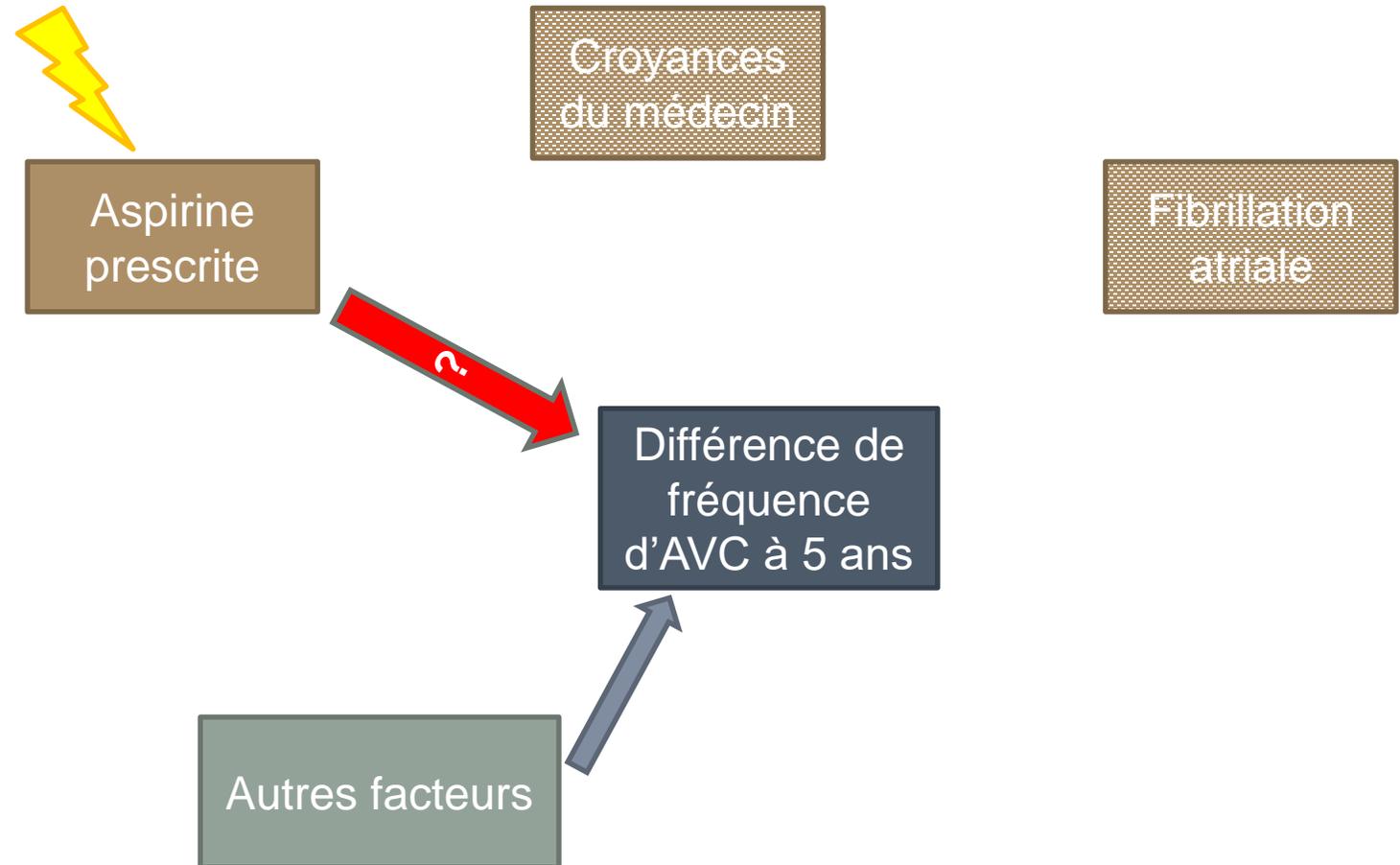
- L'allocation est rendue aléatoire : **le choix du bras (aspirine ou non) est sans cause**

Effet de la randomisation

La prescription d'aspirine est aléatoire

⇒ **Les croyances du médecin, et les autres facteurs n'ont plus d'effet sur la prescription**

• Facteurs influençant la différence



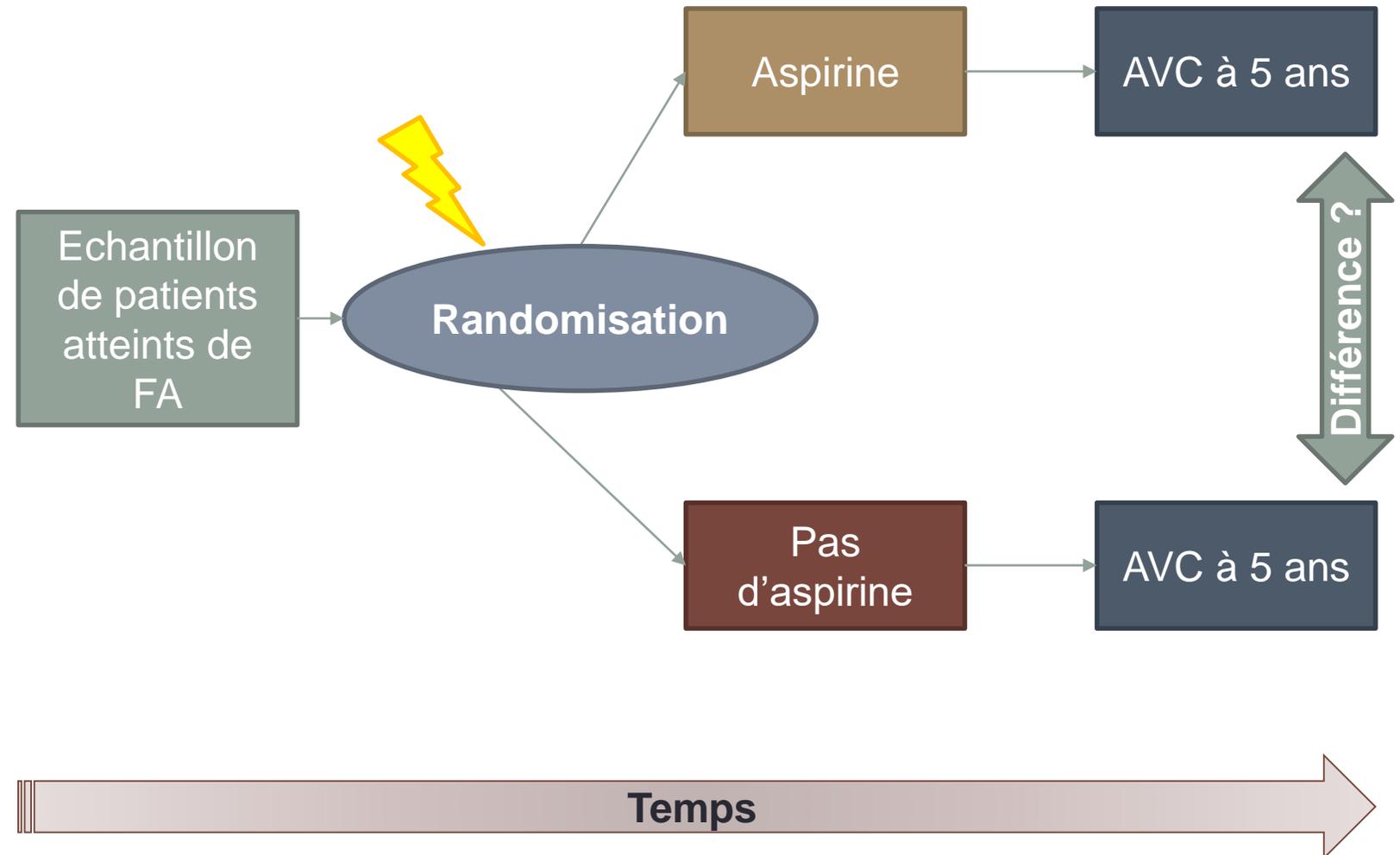
Design expérimental

En pratique, on compare deux groupes de patients entre eux :

- Avec aspirine : **intervention**
- Sans aspirine : **contrôle**

Chaque groupe est appelé un **bras**

On mesure après 5 ans de suivi la différence de fréquence d'AVC



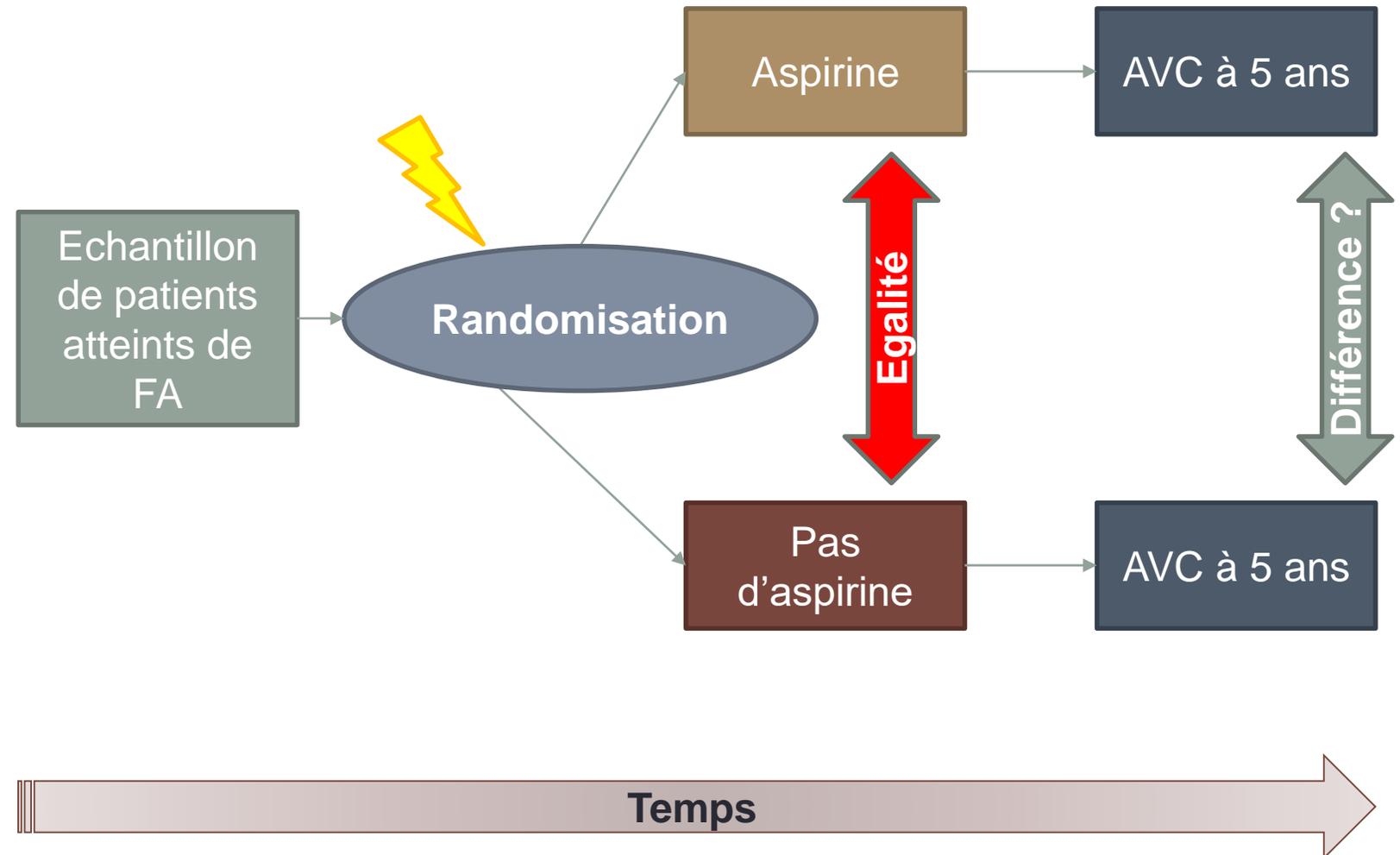
Design expérimental

Effet supplémentaire de la randomisation :

l'échantillon est distribué aléatoirement entre bras

⇒ **Les distributions des caractéristiques des deux bras sont identiques**

- ⇒ Âge
- ⇒ Sexe
- ⇒ Tabagisme
- ⇒ Diabète
- ⇒ Régime alimentaire
- ⇒ **Toute autre caractéristique non mesurée**



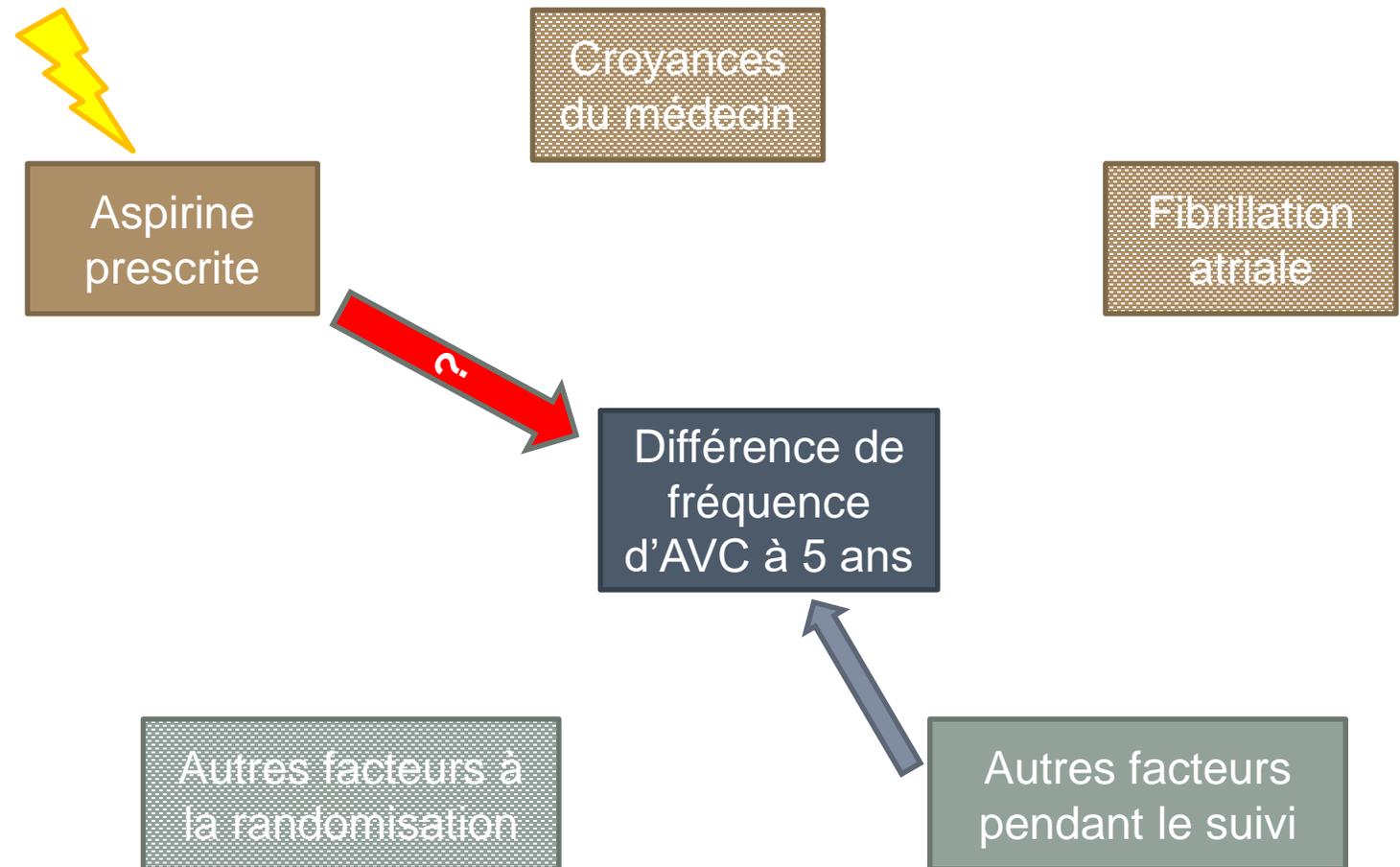
Effet de la randomisation

Au moment de la randomisation, les deux bras sont identiques (en moyenne)

⇒ **A baseline, les autres facteurs n'ont pas d'impact sur la différence de fréquence d'AVC**

Les autres facteurs peuvent évoluer différemment **après** la randomisation (devenir diabétique, débuter le tabac...)

• Facteurs influençant la différence



Définition du bras contrôle

- Plusieurs options existent pour le choix du traitement administré dans le bras contrôle :
 - Rien
 - **Placebo** : médicament factice sans effet pharmacologique
 - *Exemple : sachet de poudre inerte ayant l'apparence d'aspirine*
 - **Traitement de référence au moment de l'essai** :
 - *Exemple : essai en oncologie : usage d'un traitement prouvé comme efficace dans le bras contrôle*
- **L'éthique et la rigueur scientifique impose l'usage du meilleur traitement prouvé en tant que contrôle !**

SUIVI

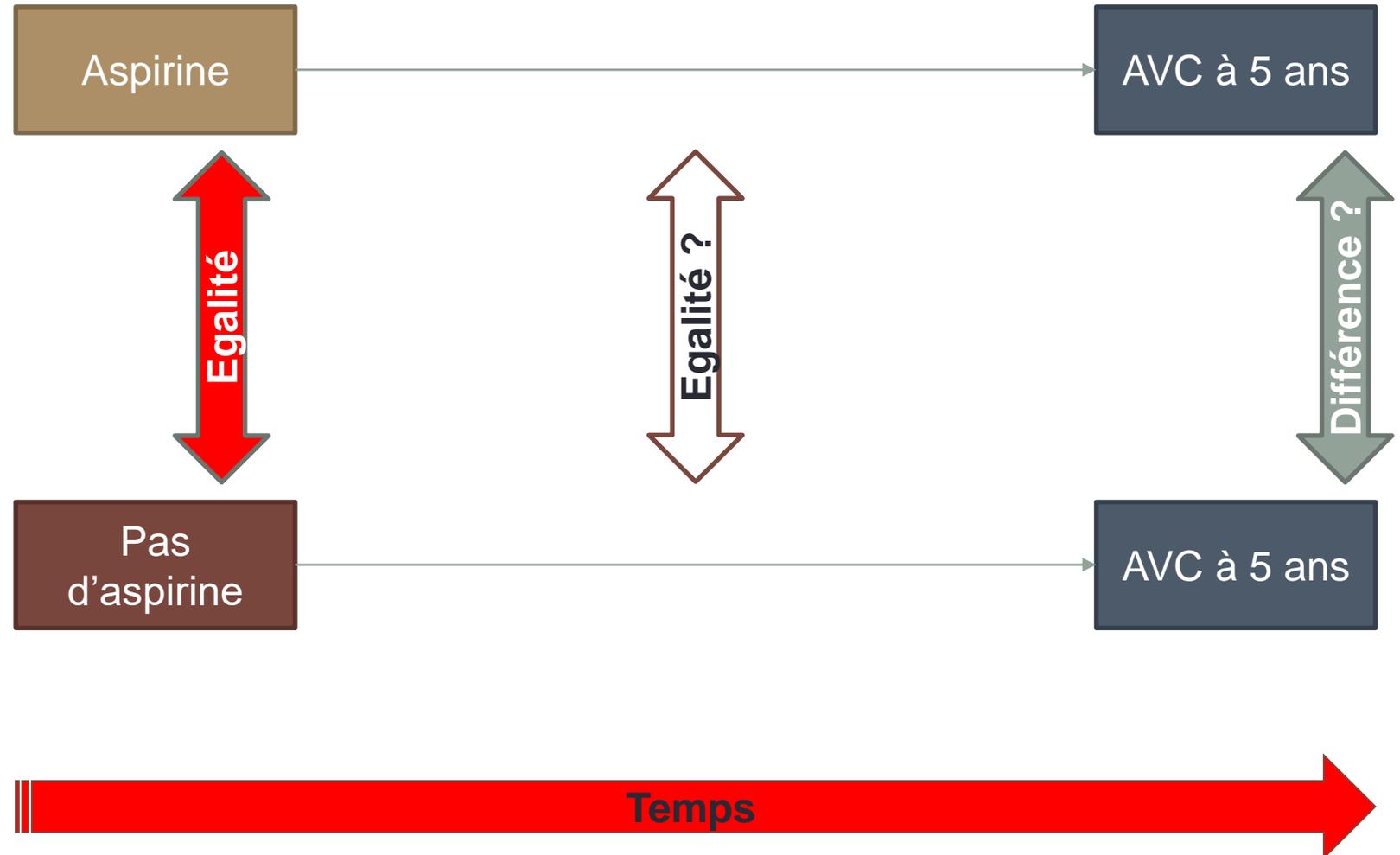
Suivi

Les traitements ne font pas effet immédiatement

⇒ Nécessité de **suivre les sujets dans le temps**

L'égalité entre bras est garantie au début par la randomisation

⇒ **Il faut garantir l'égalité pendant le suivi**



Suivi différencié

- *Exemple : un patient sous aspirine dans cet essai est suivi de façon très rapprochée par son médecin. Pendant une consultation de suivi, un diabète est découvert et immédiatement mis sous traitement.*
- Visites rapprochées > Diabète traité précocement > Réduction du risque d'AVC spécifique au bras intervention
- **Nécessité de standardiser le suivi entre groupes**

Effet placebo

- Effet d'une intervention ne se manifestant pas via le mécanisme d'action de l'intervention
- Effet psychobiologique indépendant du traitement étudié, pouvant avoir un effet mesurable sur le résultat
- *Exemple : un patient recevant de l'aspirine adapte inconsciemment son alimentation pour réduire son risque d'AVC*
- **Nécessité de prévenir l'effet placebo**

Aveugle ou insu

- **Double insu**

- Le patient et les médecins ne connaissent pas le bras d'allocation
 - ⇒ L'effet placebo est identique dans les deux bras
 - ⇒ Les soignants ne peuvent pas intervenir en fonction du bras

- **Simple insu**

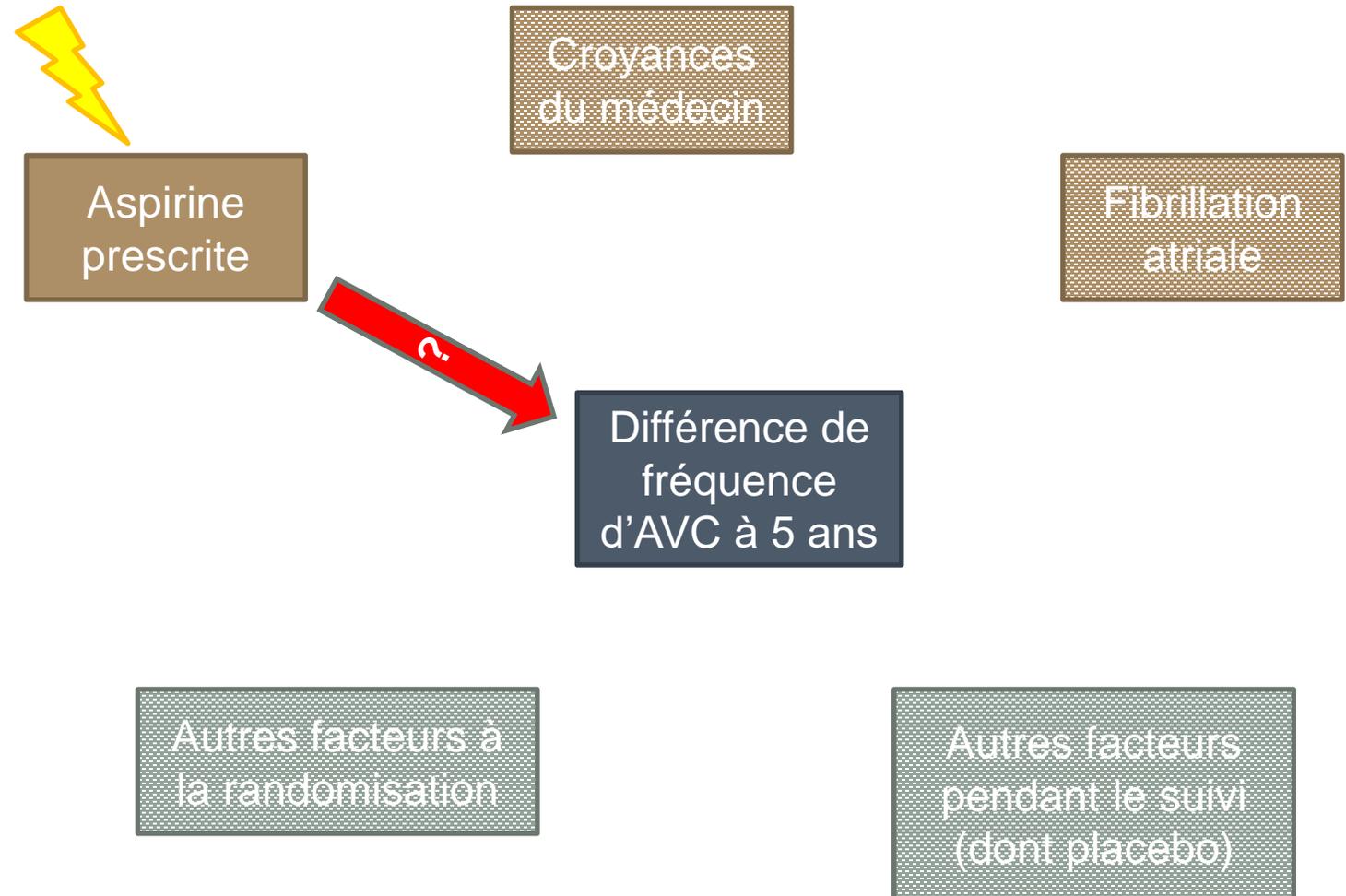
- Le patient ne connaît pas le traitement, mais le médecin oui
 - ⇒ Pas d'effet placebo différencié
 - ⇒ Nécessité de protocoliser strictement le suivi

Effet du double insu

Au cours du suivi, les deux bras sont traités identiquement

- ⇒ Pendant le suivi, les deux bras restent comparables
- ⇒ Seul le bras d'allocation est responsable d'une différence à la fin du suivi

• Facteurs influençant la différence



CRITÈRE DE JUGEMENT

Résultat d'un essai clinique

- L'efficacité de l'intervention est jugée **comparativement** au contrôle
- Le **critère de jugement** est la mesure que l'on compare entre les bras
 - *Exemple : différence de fréquence de survenue du premier AVC dans les 5 ans suivant l'inclusion*
- Ce critère doit être **objectivement mesurable**
 - *Exemple : diagnostic d'AVC sur des critères internationalement reconnus, par des neurologues indépendants de l'étude*

Critères de jugement

Principal

- Focus de l'analyse
- Utilisé pour juger de l'efficacité du traitement
- **Utilisé pour formuler les recommandations de traitement**

Secondaires

- Information complémentaire
- Non utilisé dans les recommandations
- Particularité : effets indésirables
 - Est-ce que la nouvelle intervention est plus dangereuse que le contrôle ?

Exemples de critères de jugement

- Décès
- Hospitalisation
- Progression d'une maladie (*cancer*)
- Survenue d'une maladie (*infarctus du myocarde*)
- Questionnaire de qualité de vie (*échelle SF36*)
- Dosage sanguin (*HbA1c en diabète*)
- Effets indésirables (*toxidermie*)
- Epreuve fonctionnelle (*fonction respiratoire, mobilité d'une articulation...*)

ANALYSE

Analyse du résultat

- A l'issue de l'essai :
 - 1 mesure du CJP dans le bras intervention par patient
 - 1 mesure du CJP dans le bras contrôle par patient
- Comparaison du CJP entre les bras par un test d'hypothèse
 - Rejeter H_0 conduit à retenir la supériorité d'un traitement sur l'autre
 - **Effet du traitement** : différence/ratio du CJP entre les bras
 - *Exemple : différence de fréquence de premier AVC dans les 5 ans suivants l'inclusion : 12%, IC 95% [9, 15], p-value=0,006*

Analyse en Intention de Traiter (ITT)

- L'analyse principale de l'essai respecte deux conditions
 - L'intégralité des patients **randomisés** sont analysés
 - Les patients sont analysés **dans le bras alloué par la randomisation**
- Toute autre analyse **brise la randomisation** et ses bénéfices
- En pratique, durant un essai :
 - Les patients ne sont plus suivis de façon prématurée : perte de vue, retrait de consentement > **Données manquantes**
 - Les patients **changent de bras** (*en oncologie : progression dans le bras contrôle > passage au traitement du bras intervention*)
 - Les patients ne sont pas traités en accord avec le protocole (*ajustement de la posologie, ajout d'un traitement non prévu...*)

Analyse en Intention de Traiter (ITT)

- **Données manquantes** : le CJP n'est pas disponible
 - **Remplacer** par des valeurs probables
 - ⇒ **Perte de précision** (IC plus larges, p-value plus élevée)
- **Changement de bras** : le patient n'a pas reçu uniquement le traitement alloué
 - **Analyse comme si** le traitement alloué était administré
 - ⇒ **Diminution de la taille d'effet** (différence rapprochées de 0, ratios rapprochés de 1)
- **L'analyse en ITT est conservatrice** : diminue la probabilité de déclarer la supériorité à tort comparativement aux autres modes d'analyse

ESSAI CLINIQUE CONTRÔLÉ RANDOMISÉ

Gold standard de l'évaluation thérapeutique

- **Essai** : expérience planifiée en amont des faits
 - **Contrôlé** : présence du bras contrôle, auquel on compare le bras intervention
 - **Randomisé** : le choix du bras (intervention ou contrôle) est aléatoire
 - **En double aveugle** : pas de différence dans l'effet placebo, ni de prise en charge différenciée
 - **Critère de jugement objectif** : mesure peu ou pas dépendante du médecin
- ⇒ **Statistiquement équivalent à remonter le temps pour tester deux fois un même patient**

Essai / expérience

Essai - Expérimentation

- L'investigateur décide du déroulé de l'essai
- On maîtrise :
 - Les critères d'inclusion et d'exclusion
 - Le protocole
 - Le nombre de sujets
 - Les traitements évalués (intervention et contrôle)
 - Les visites de suivi
 - Les critères de jugement

Etude - Observation

- L'investigateur analyse des dossiers médicaux
 - Pas de vérification des critères d'inclusion et d'exclusion
 - Pas de randomisation
 - Pas d'insu
 - Pas de suivi standardisé
 - Pas de critère de jugement standardisé

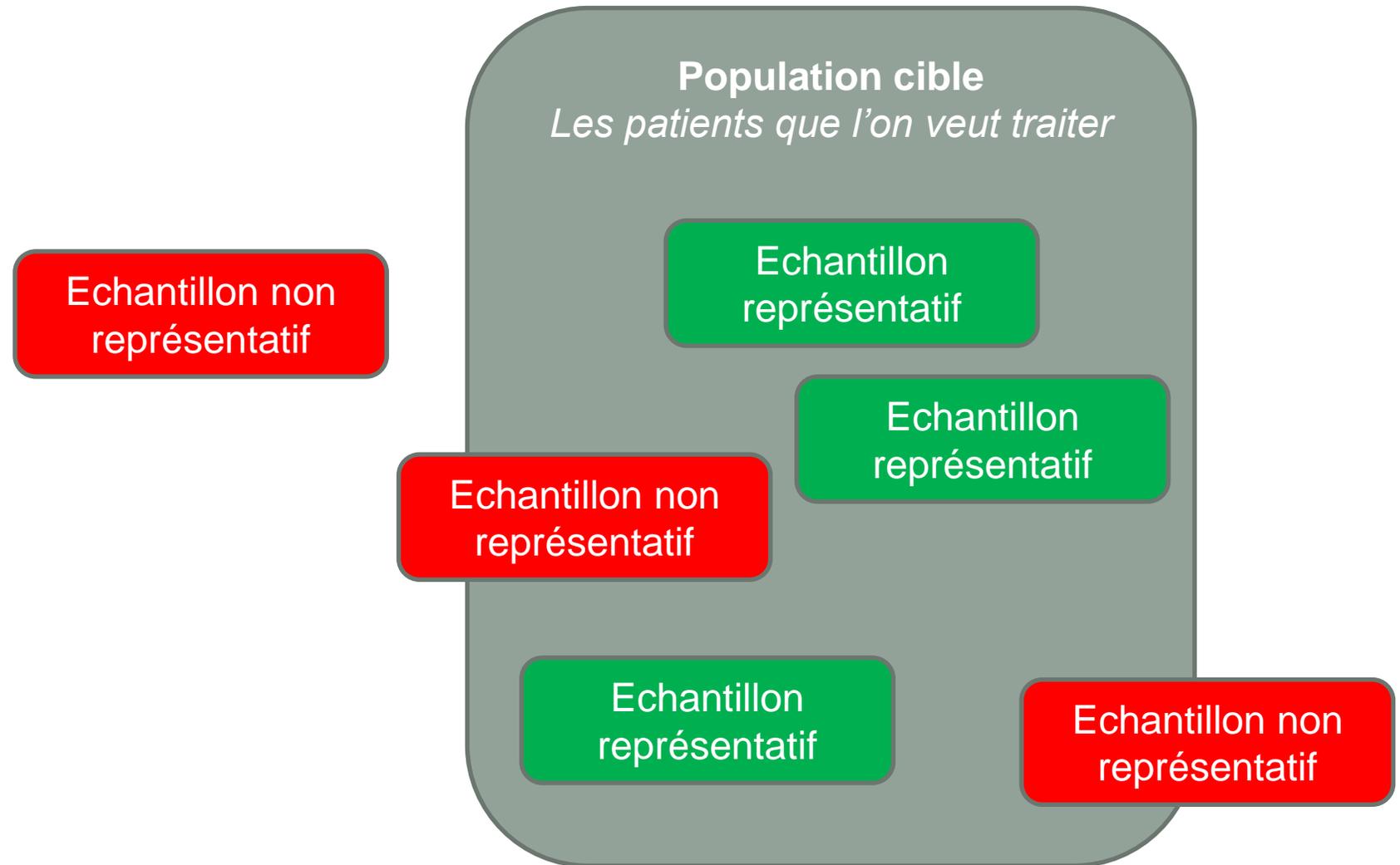
Biais de sélection

Déviations entre la population cible de l'étude (les personnes que l'on veut traiter) et l'échantillon

L'échantillon **n'est plus représentatif** de la population cible

- Mauvaise définition des critères d'inclusion et d'exclusion

⇒ **Le résultat de l'essai n'est pas applicable à la population cible**



Essai comparatif / monobras

Contrôlé - Comparatif

- On compare le traitement étudié à un traitement de référence administré dans les mêmes conditions

⇒ **Garantie de comparabilité**

Pas de contrôle - Monobras

- On compare le traitement étudié à des données tierces (littérature scientifique, autre centre hospitalier...)

⇒ **Aucune garantie de comparabilité** (populations différentes entre les bras)

Randomisé / Non randomisé

Randomisé

- Aucun facteur n'influence la décision de traitement
- ⇒ **Absence de confusion** entre l'effet du traitement avec les facteurs influençant la prescription

Non randomisé

- Le médecin choisit de prescrire ou non le traitement
- ⇒ **Confusion** entre biais du médecin et l'effet du nouveau traitement

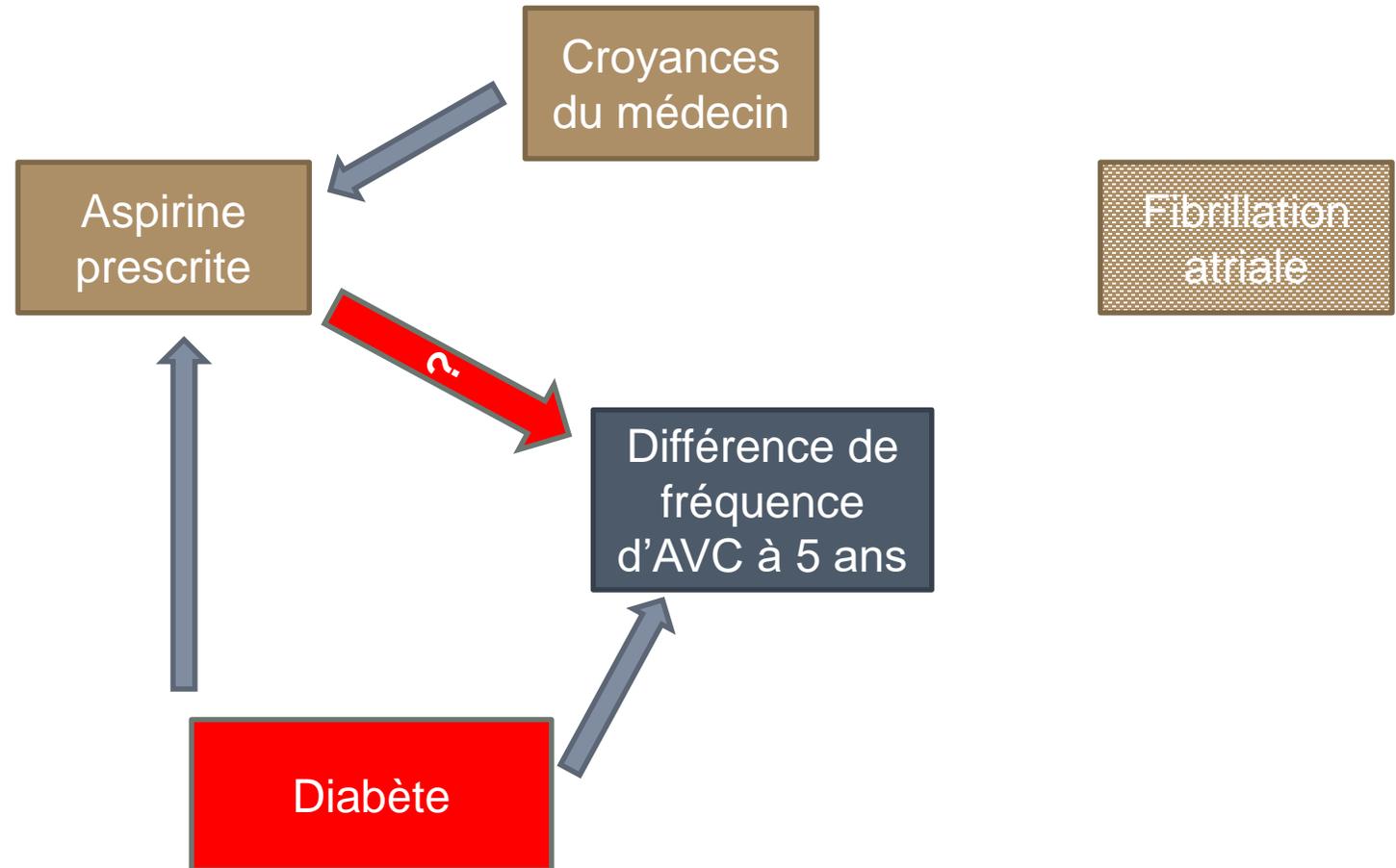
Biais de confusion

Un biais de confusion existe lorsque des **facteurs influencent le choix du bras et le résultat de l'étude**

- Sans randomisation, le diabète modifie la probabilité de prescription et également la probabilité d'AVC

⇒ Le diabète est un **facteur de confusion**

• Facteurs influençant la différence



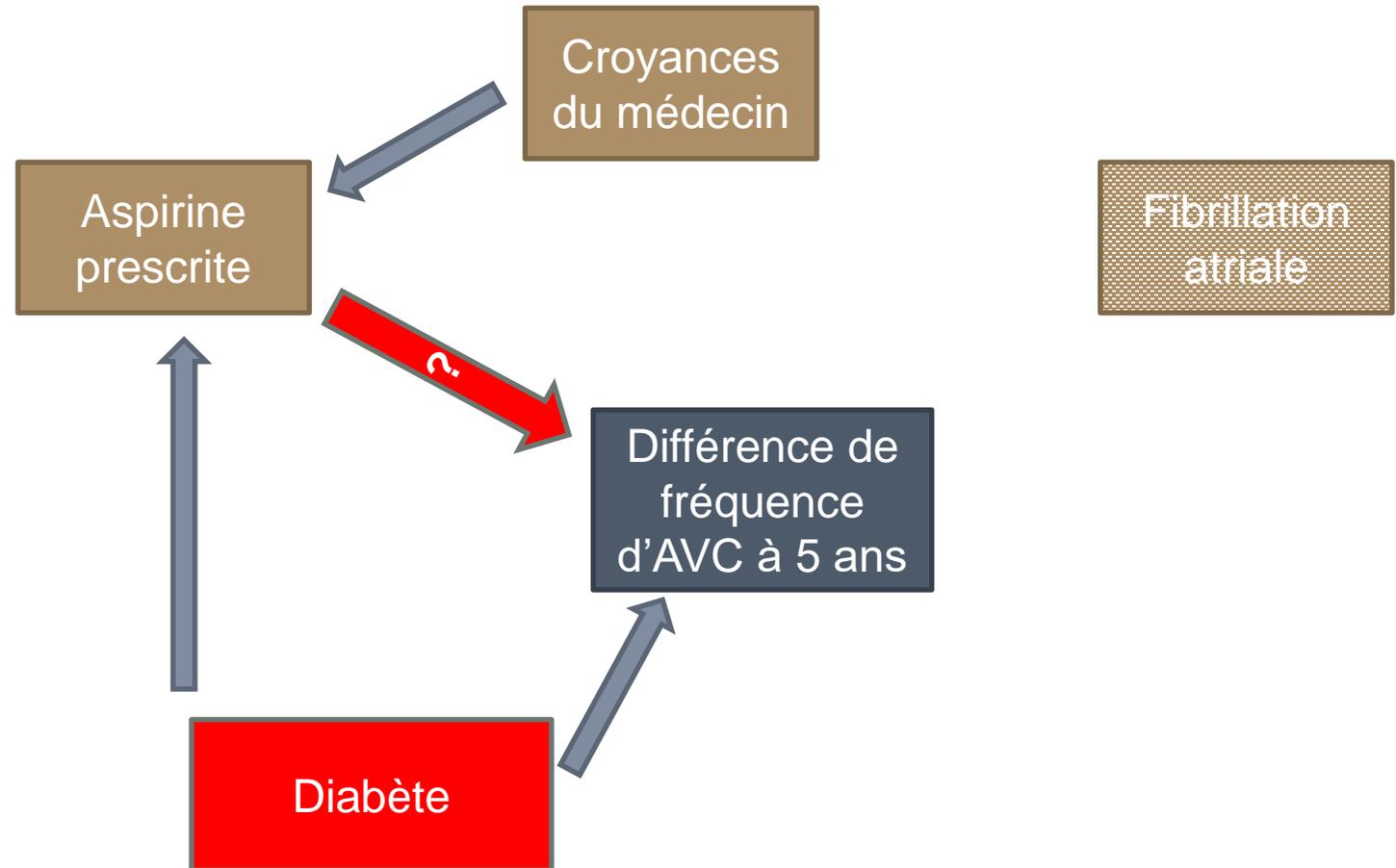
Biais de confusion

La différence de fréquence d'AVC à 5 ans en présence de confusion est **attribuable à l'aspirine ET au diabète** de façon combinée

On ne peut pas séparer ces deux effets sans randomisation

La randomisation empêche tout biais de confusion à baseline

• Facteurs influençant la différence



Biais de confusion

La randomisation est la seule méthode prévenant totalement le biais de confusion à baseline.

Insu / Ouvert

Double insu – Bras inconnu

- Suivi identique entre les bras
⇒ Pas de **biais de suivi**
- Mesure identique du critère de jugement
⇒ Pas de **biais de mesure**

Ouvert – Bras connu

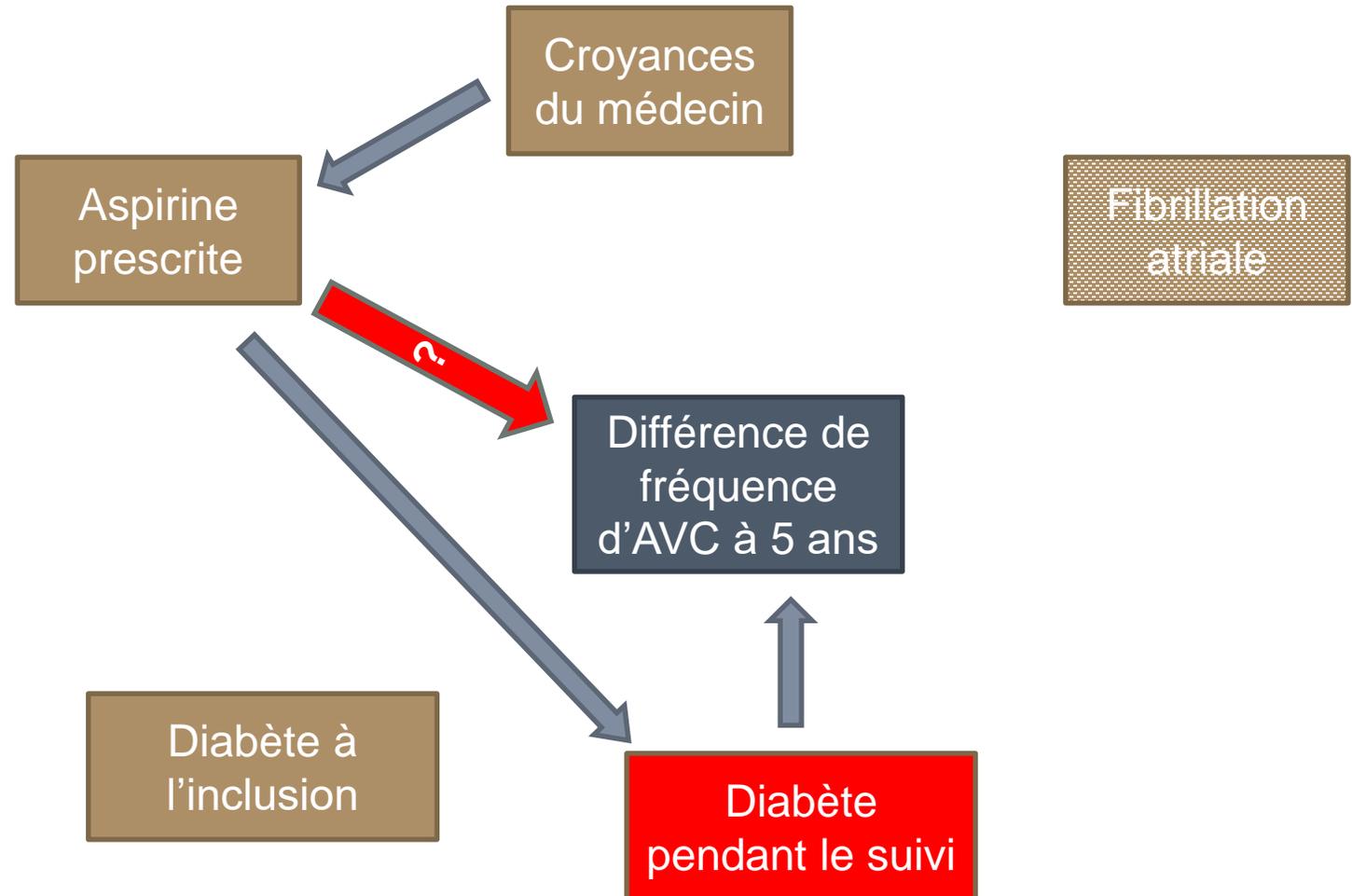
- Suivi potentiellement différencié
⇒ Possible **biais de suivi**
- Mesure différenciée du critère
⇒ Possible **biais de mesure**
⇒ Majoré si critère peu objectif

Aveugle et biais de suivi

En absence d'aveugle, on peut introduire un **bias de suivi**

- L'aspirine augmente la probabilité de traiter un diabète dans le bras intervention
 - Les patients sont mieux traités pour leur diabète dans le bras intervention
- ⇒ La différence mesurée correspond à **l'effet de l'aspirine ET l'effet d'un suivi plus intense**

• Facteurs influençant la différence



Biais de mesure

- Evaluation du critère de jugement en connaissance du bras d'allocation du patient
- **Biais du clinicien en faveur d'un groupe ou l'autre**
- **L'aveugle garantit l'absence de biais de mesure**
- Une évaluation d'un critère de jugement standardisé par un comité indépendant en insu du bras d'allocation compense l'absence d'insu

PUISSANCE STATISTIQUE ET ESSAI CLINIQUE

Nombre de sujets nécessaires

- Essai clinique => **étude expérimentale**
 - On fait un essai pour **prouver une taille donnée de traitement** (hypothèse de recherche)
 - On contrôle le déroulé de l'étude, notamment la taille de **l'échantillon**
- ⇒ On peut calculer un **nombre de sujets nécessaires** pour une **hypothèse de taille d'effet**

Rappel

Si l'hypothèse nulle est fautive, $1 - \beta$ est la puissance du test d'hypothèse.

		L'hypothèse nulle (H_0) est...	
		Vraie	Fausse
Décision sur l'hypothèse nulle (H_0)	Ne pas rejeter	Inférence correcte (vrai négatif) (probabilité = $1 - \alpha$)	Erreur de type II (faux négatif) (probabilité = β)
	Rejeter	Erreur de type I (faux positif) (probabilité = α)	Inférence correcte (vrai positif) (probabilité = $1 - \beta$)

Puissance, $1 - \beta$

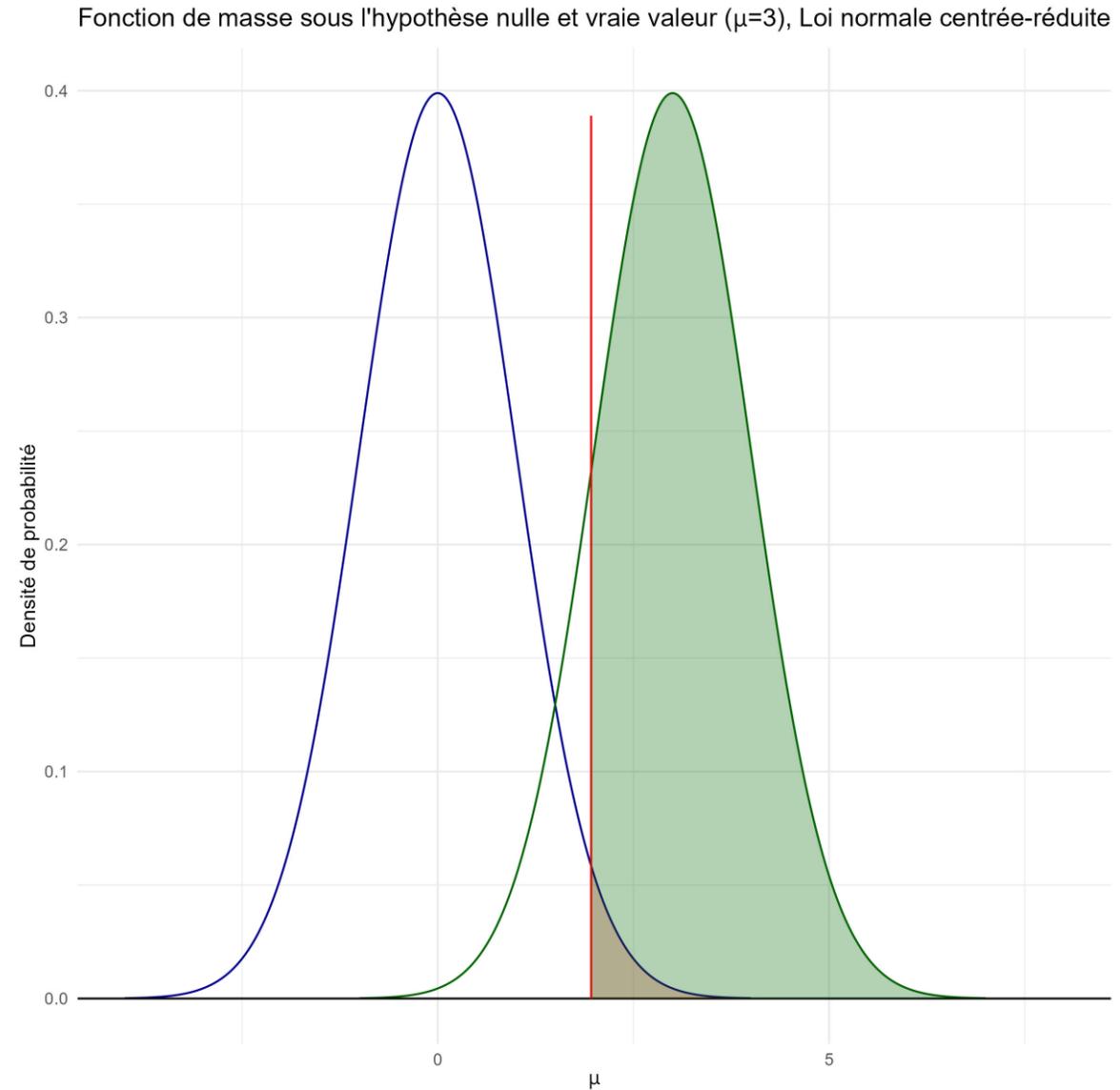
- La **puissance**, complément du risque de seconde espèce, « mesure » la capacité d'un test à détecter une **différence significative spécifique** (pour une « vraie valeur » donnée).
- Elle est utilisée en **calcul du nombre d'effectifs** : « De combien de sujets ai-je besoin dans mon expérience pour prouver une différence avec une puissance de 80% ? »

Puissance

La puissance est le complément de β , donc **l'aire sous la vraie distribution définie par la gauche** par le seuil de rejet.

La puissance varie en fonction de :

- α
- $\mu_0 - \mu$
- σ
- n via l'erreur-type

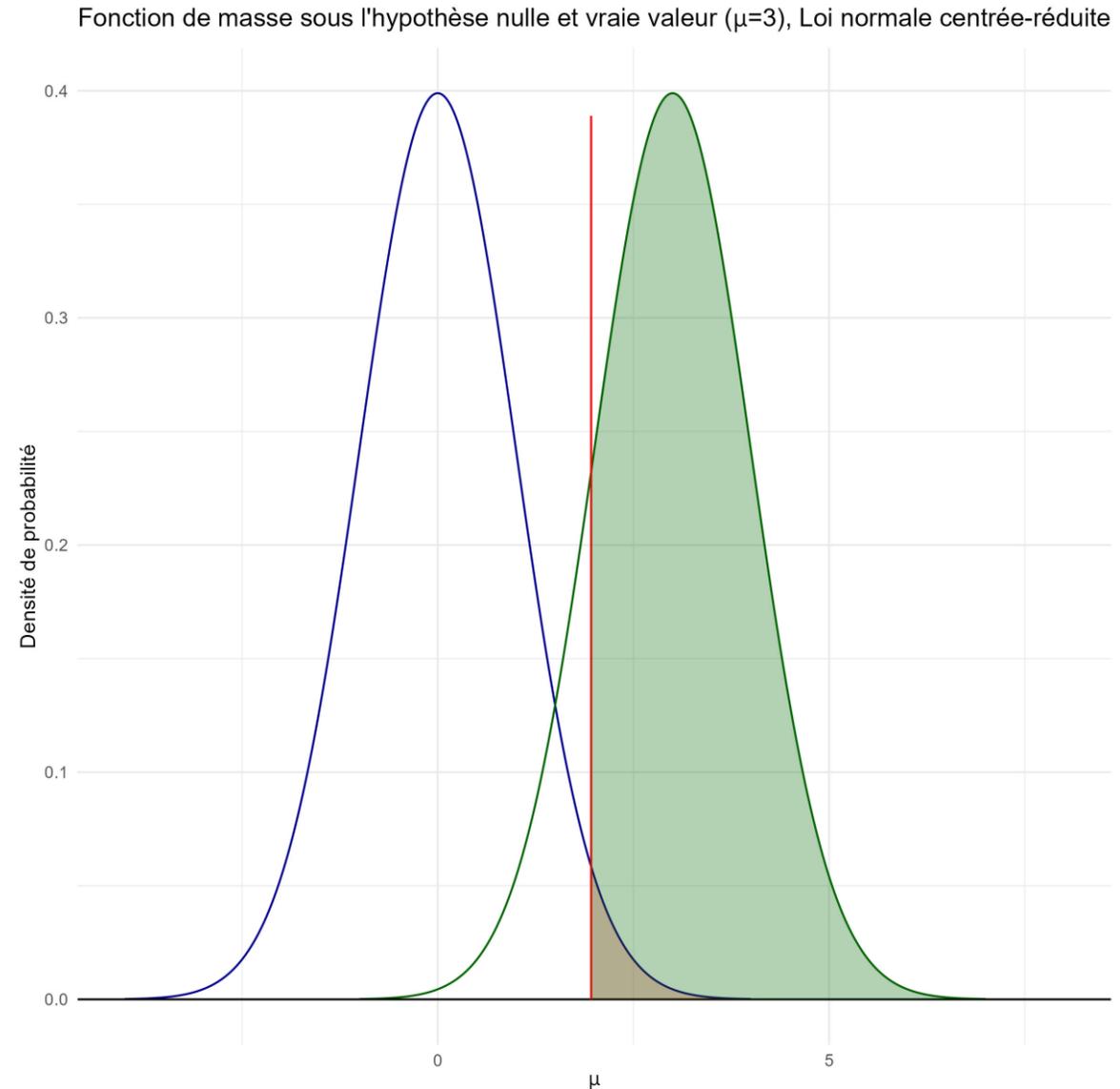


Puissance

Connaître la puissance requiert de connaître la vraie valeur du paramètre.

Donc on ne connaît jamais la puissance effective d'un test.

Le seul moment où la puissance est connue avec certitude, c'est lors du calcul du nombre de sujets nécessaire, où l'on présume une « vraie valeur ».



Prérequis au calcul du NSN

Pour l'exemple de la différence de risque d'AVC à 5 ans :

- Hypothèse de probabilité d'événement dans le groupe contrôle :
 $\pi_C = \pi_0 = 0,10$
- Hypothèse de probabilité d'événement dans le groupe intervention : $\pi_I = 0,05$
- Hypothèse de différence de probabilité :
$$\delta = \pi_C - \pi_I = 0,1 - 0,05$$
- Risque $\alpha = 0,05$: **choisi par l'investigateur**
- Puissance $1 - \beta = 0,8$: **choisi par l'investigateur**

Hypothèses du calcul du NSN

- δ : hypothèse d'effet, H_a , une des valeurs possibles dans H_1
- Ceci n'est qu'une **hypothèse**, définie **avant la réalisation de l'essai**
- Cette hypothèse peut être invalidée par l'essai (traitement plus ou moins efficace qu'attendu)
- Cette hypothèse n'est qu'une seule des possibilités dans H_1

Exemple de calcul

- Pour une simple différence de proportions, le NSN pour un essai de supériorité est calculable directement :

$$n = \frac{(z_{1-\beta} + z_{1-\alpha/2})^2 \pi_I(1 - \pi_I) + \pi_C(1 - \pi_I)}{\delta^2}$$

Avec :

- $z_{1-\beta}$ quantile de la Loi Normale centrée réduite au niveau $1 - \beta$
- $z_{1-\alpha/2}$ quantile de la LNCR au niveau $1 - \alpha/2$
- *Exemple illustratif, la réalisation du calcul n'est pas au programme*

SYNTHÈSE

Les piliers méthodologique de l'essai clinique

Critères d'inclusion et d'exclusion

- Correspondre à la population cible
- **Prévenir le biais de sélection**

Définition du bras contrôle

- Selon les recommandations en vigueur
- Ethique

Randomisation

- **Prévenir le biais de confusion**

Double insu

- **Prévenir le biais de suivi et le biais de mesure**

CJP objectif et reconnu

- **Prévenir le biais de mesure**
- Comparer au reste de la littérature

Analyse en Intention de Traiter (ITT)

- Analyse conservatrice : prévenir les résultats faussement significatifs, respecter la randomisation

A retenir

- L'essai clinique est une expérience planifiée pour répondre à une question spécifique
- Le respect des piliers méthodologiques garantit de mesurer l'effet du nouveau traitement comparé à l'ancien
- Le choix du bras contrôle doit être adapté à la pathologie et les recommandations en vigueur
- Le calcul du NSN dépend d'une hypothèse de taille d'effet, qui n'est qu'une valeur possible de H_1

MERCI POUR VOTRE
ATTENTION
