

Poly de révision

Analyses intermédiaires 2026, alpha-spending et réponses d'examen

Table des matières

Poly de révision	2
Analyses intermédiaires 2026, alpha-spending et réponses d'examen	2
Comment utiliser ce poly	2
1. Carte rapide des questions couvertes	2
2. Reconnaître une vraie analyse intermédiaire	3
2.1. Définition	3
2.2. Trois grands objectifs	3
2.3. Pourquoi on ne peut pas "regarder quand on veut"	3
3. Les briques de base à connaître	3
3.1. Le vocabulaire minimal	3
3.2. Efficacité, futilité, sécurité	4
3.3. O'Brien-Fleming, Pocock, Lan-DeMets	4
3.4. Bilatéral, unilatéral, borne effective	4
3.5. Pourquoi un effet vu tôt peut être surestimé	5
3.6. Quand le raisonnement ne s'applique pas	5
4. Pièges fréquents	5
5. Comment écrire une bonne réponse d'examen	5
6. Articles résolus avec questions fictives	6
6.1. 2026 RC	6
Article et contexte	6
Questions fictives proposées	6
Extrait de l'article à analyser	6
Ce que l'examineur veut vérifier	6
Corrections détaillées	7
6.2. 2026 EPI	9
Article et contexte	9
Questions fictives proposées	9
Extrait de l'article à analyser	10
Ce que l'examineur veut vérifier	10
Corrections détaillées	10
7. Réponses prêtes à l'emploi	12
7.1. Si l'on vous demande de définir une analyse intermédiaire	12
7.2. Si l'on vous demande pourquoi il faut contrôler l'alpha	13
7.3. Si l'on vous demande le rôle du DSMB / IDMC	13
7.4. Si l'on vous demande pourquoi ce raisonnement ne s'applique pas à une cohorte naturaliste	13
8. Checklist finale	13
9. Conclusion	13

Poly de révision

Analyses intermédiaires 2026, alpha-spending et réponses d'examen

Document de travail orienté examen : définitions, pièges, stratégie de dépense alpha, et vingt questions fictives détaillées à partir des articles RC 2026 et EPI 2026.

Ce document reprend la logique des polys d'annales existants, mais il est entièrement centré sur les analyses intermédiaires, la dépense du risque alpha, le rôle du monitoring indépendant, et la manière d'écrire une réponse complète d'examen sur ce thème.

Ce poly a un objectif simple : **permettre de traiter correctement les questions d'examen portant sur les analyses intermédiaires.**

Il suit la même logique que les autres polys du dossier :

1. comprendre exactement ce qu'est une analyse intermédiaire ;
2. distinguer efficacité, futilité et sécurité ;
3. reconnaître quand ce raisonnement a du sens, et quand il n'en a pas ;
4. répondre année/article par année/article avec des formulations directement réutilisables.

analyses intermédiaires alpha-spending DSMB / IDMC annales fictives HTML A4 + PDF

! Idée directrice

Une **analyse intermédiaire** n'est pas une simple lecture précoce de résultats.

C'est une **analyse planifiée**, intégrée au protocole, avec des règles explicites pour préserver l'interprétation statistique et clinique de l'essai.

Et surtout : **ce cadre a du sens dans un essai prospectif**, pas dans une étude observationnelle naturaliste rétrospective.

Comment utiliser ce poly

Lecture conseillée :

1. lire d'abord les sections 1 à 5 pour maîtriser la mécanique générale ;
2. travailler ensuite les deux grandes sections 2026 RC puis 2026 EPI ;
3. finir par les réponses prêtes à l'emploi et la checklist.

Si vous êtes en révision rapide, les éléments à mémoriser en priorité sont :

- ce qu'est une vraie analyse intermédiaire ;
- la différence entre arrêt pour efficacité, futilité et toxicité ;
- l'idée générale de O'Brien-Fleming, Pocock et Lan-DeMets ;
- la manière d'expliquer pourquoi ce raisonnement est **approprié dans un essai** mais **inadapté dans une cohorte naturaliste**.

1. Carte rapide des questions couvertes

Article	Type de design	Nature des questions	Mouvement attendu
2026 RC	essai randomisé phase 3	vraie logique d'analyses intermédiaires	expliquer, commenter et interpréter la stratégie statistique
2026 EPI	étude observationnelle naturaliste	absence de vraie logique d'interim + transposition hypothétique	dire pourquoi le cadre est différent, puis raisonner sur un protocole fictif

2. Reconnaître une vraie analyse intermédiaire

2.1. Définition

Une analyse intermédiaire est une analyse **planifiée avant la fin de l'essai**, inscrite dans le protocole ou le SAP, déclenchée à un nombre d'événements, à une date, ou après un certain volume de suivi.

Elle n'est pas synonyme de :

- regard exploratoire sur les données ;
- monitoring descriptif de routine ;
- sous-analyse opportuniste ;
- compte-rendu préliminaire sans règle de décision.

2.2. Trois grands objectifs

Une analyse intermédiaire peut viser :

- **l'efficacité** : arrêter tôt si le bénéfice est déjà évident ;
- **la futilité** : arrêter si l'essai a peu de chances d'atteindre son objectif ;
- **la sécurité** : protéger les participants en cas de signal de toxicité.

2.3. Pourquoi on ne peut pas “regarder quand on veut”

Si l'on teste plusieurs fois sans correction, on augmente artificiellement la probabilité de faux positif.

En version très schématique :

$$\alpha_{\text{global}} \neq \alpha_1 + \alpha_2 + \alpha_3$$

mais **plus on multiplie les looks naïfs, plus le risque de première espèce réel dérive au-dessus du seuil annoncé.**

La fonction d'alpha-spending sert précisément à éviter cela.

3. Les briques de base à connaître

3.1. Le vocabulaire minimal

Élément	Ce qu'il faut savoir dire
analyse intermédiaire	analyse planifiée avant l'analyse finale
borne	seuil critique à franchir pour une décision
alpha-spending	manière de "dépenser" progressivement le risque alpha
IDMC / DSMB	comité indépendant chargé d'examiner les données de sécurité et d'efficacité intermédiaires
futilité	faible probabilité de succès final même si l'essai continue
data-cutoff	date ou moment auquel les données sont gelées pour l'analyse

3.2. Efficacité, futilité, sécurité

Une bonne copie distingue toujours ces trois objectifs :

- l'arrêt pour **efficacité** dit : "le bénéfice paraît déjà suffisamment démontré" ;
- l'arrêt pour **futilité** dit : "continuer a peu de chances de modifier la conclusion" ;
- l'arrêt pour **sécurité** dit : "continuer pourrait exposer inutilement les participants".

Ces trois logiques ne reposent pas forcément sur les mêmes seuils ni sur les mêmes critères.

3.3. O'Brien-Fleming, Pocock, Lan-DeMets

Méthode	Idée simple	Ce qu'il faut écrire à l'examen
O'Brien-Fleming	très conservatrice tôt, moins conservatrice tard	protège fortement contre un faux positif précoce
Pocock	seuil plus constant d'un look à l'autre	plus simple à comprendre, mais dépense plus vite l'alpha
Lan-DeMets	version flexible d'une logique d'alpha-spending	permet de s'adapter aux looks réellement observés

Formulation à retenir :

Une frontière O'Brien-Fleming rend les premières analyses très exigeantes. Une approche Lan-DeMets permet d'implémenter cette idée de manière souple sans figer exactement le calendrier initial.

3.4. Bilatéral, unilatéral, borne effective

Dans les essais, on voit souvent deux notations :

$$z_{1-\alpha/2}$$

pour un test bilatéral, et

$$z_{1-\alpha}$$

pour un test unilatéral.

Dans un plan avec analyses intermédiaires, on ne réutilise pas simplement le seuil final standard à chaque look. On obtient un **alpha effectif** plus ou moins strict selon le moment de l'analyse.

3.5. Pourquoi un effet vu tôt peut être surestimé

Un résultat très précoce peut :

- capter une fluctuation favorable temporaire ;
- reposer sur des données encore immatures ;
- surestimer la vraie taille d'effet à long terme.

Cela ne veut pas dire qu'un arrêt précoce est "faux", mais qu'il doit être discuté avec prudence.

3.6. Quand le raisonnement ne s'applique pas

Dans une étude observationnelle rétrospective naturaliste :

- il n'y a pas de randomisation prospective à protéger ;
- il n'y a pas en général de protocole de looks successifs avec bornes formelles ;
- le vrai enjeu est plus souvent la **validité causale**, le **biais**, la **confusion** et la **qualité des données**.

Dire cela clairement rapporte des points.

Encadré alpha-spending

L'idée centrale de l'alpha-spending est de **préserver le risque global de première espèce** malgré plusieurs analyses successives.

On "dépense" donc l'alpha petit à petit, avec des seuils souvent très stricts au début et moins stricts à la fin.

Formulation type à réutiliser

Une analyse intermédiaire n'est acceptable que si elle était prévue, que les règles de décision sont explicites, et que le contrôle du risque alpha global est assuré. Sinon, on transforme un essai confirmatoire en suite de regards opportunistes sur les données.

4. Pièges fréquents

- confondre monitoring descriptif et vraie analyse intermédiaire ;
- croire qu'une simple lecture à mi-parcours suffit sans correction statistique ;
- oublier que le rôle du DSMB / IDMC est indépendant de l'équipe d'investigation ;
- croire qu'un résultat très précoce est automatiquement plus convaincant ;
- parler d'analyses intermédiaires confirmatoires dans une étude observationnelle rétrospective comme s'il s'agissait d'un essai randomisé.

5. Comment écrire une bonne réponse d'examen

La structure qui fonctionne presque toujours est :

1. rappeler de quel type d'analyse il s'agit ;

2. dire pourquoi elle était pertinente ou non ;
3. nommer explicitement la méthode de contrôle du risque alpha ;
4. commenter la conséquence clinique et méthodologique ;
5. finir par une conclusion nuancée.

Phrase-type :

Il s'agit d'une analyse intermédiaire planifiée, intégrée au protocole, avec contrôle du risque alpha par une stratégie de type [méthode]. Elle est donc interprétable, mais ses résultats doivent être commentés à la lumière de la maturité encore incomplète des données.

6. Articles résolus avec questions fictives

6.1. 2026 RC

Article et contexte

Article support : **Fang et al., NEJM 2026**, essai de phase 3 OptiTR0P–Lung04 comparant **sacituzumab tirumotecan** à une chimiothérapie standard dans le NSCLC avancé muté EGFR résistant aux TKI.

- essai randomisé 1:1, multicentrique, ouvert ;
- 376 patients randomisés ;
- critère principal : progression-free survival (PFS) évaluée par BICR ;
- critère secondaire hiérarchisé majeur : overall survival (OS) ;
- vraie logique d'analyses intermédiaires intégrée au plan statistique ;
- matériau particulièrement adapté à des questions d'examen sur Lan-DeMets, O'Brien-Fleming, IDMC et interprétation des seuils.

Questions fictives proposées

1. Qu'est-ce qu'une analyse intermédiaire dans le contexte de cet essai ?
2. Pourquoi une analyse intermédiaire de PFS était-elle pertinente ici ?
3. Expliquez la logique Lan-DeMets / O'Brien-Fleming.
4. Commentez le seuil alpha effectif retenu à l'interim de PFS.
5. Quel est le rôle exact du IDMC dans cet essai ?
6. Discutez l'intérêt d'une analyse intermédiaire de OS.
7. Pourquoi la borne de OS est-elle plus stricte ?
8. Que peut-on conclure si une borne d'interim est franchie ?
9. Pourquoi un résultat précoce peut-il surestimer la taille d'effet ?
10. Rédigez une réponse d'examen complète sur la stratégie d'analyses intermédiaires de cet essai.

Extrait de l'article à analyser

Ce que l'examineur veut vérifier

- savez-vous reconnaître une **vraie analyse intermédiaire planifiée** ;
- savez-vous expliquer pourquoi on ne peut pas réutiliser un seuil standard à chaque look ;
- savez-vous distinguer PFS, OS, rôle du IDMC et contrôle du risque alpha ;
- savez-vous écrire une réponse clinique **et** statistique, pas seulement réciter des mots-clés.

Corrections détaillées

Question 1

Qu'est-ce qu'une analyse intermédiaire dans le contexte de cet essai ?

Réponse type

Dans cet essai, une analyse intermédiaire est une **analyse d'efficacité planifiée avant l'analyse finale**, déclenchée ici sur un nombre minimal d'événements de PFS. Elle n'est pas opportuniste : elle est intégrée au protocole, reliée à une fonction d' α -spending, et interprétée sous le contrôle d'un comité indépendant. La bonne définition d'examen est donc : **une lecture précoce, mais prévue, encadrée statistiquement et décisionnellement.**

Question 2

Pourquoi une analyse intermédiaire de PFS était-elle pertinente ici ?

Réponse type

Elle était pertinente parce que le PFS est le critère principal, qu'il survient plus tôt que l'OS, et qu'il permet donc de détecter plus rapidement un signal d'efficacité. En oncologie avancée, attendre systématiquement l'analyse finale complète peut retarder l'accès à un traitement potentiellement utile. Une interim PFS bien encadrée permet donc d'allier **rapidité de décision** et **rigueur statistique**. Il faut toutefois rappeler que cette rapidité justifie justement l'existence d'une borne plus exigeante qu'un simple seuil habituel.

Question 3

Expliquez la logique Lan-DeMets / O'Brien-Fleming.

Réponse type

La logique O'Brien-Fleming consiste à être **très conservateur lors des premiers looks** puis à devenir moins strict à mesure que l'essai se rapproche de sa fin. La méthode Lan-DeMets permet d'implémenter cette logique de façon flexible, même si le calendrier réel des analyses diffère légèrement de celui imaginé au protocole. L'idée essentielle à formuler en copie est la suivante : **on ne dépense pas tout l'alpha d'un coup ; on le répartit progressivement pour protéger le risque global de première espèce.**

Question 4

Commentez le seuil alpha effectif retenu à l'interim de PFS.

Réponse type

Le papier indique un alpha effectif de $0,0337$ bilatéral pour cette interim PFS. Ce chiffre n'est pas un détail décoratif : il montre qu'on ne teste pas "comme d'habitude", mais selon une borne adaptée au fait qu'il s'agit d'une lecture anticipée. Le bon commentaire est donc : **le seuil utilisé à ce look est explicitement recalibré pour préserver l'alpha global de l'essai.** L'examineur attend surtout que vous disiez pourquoi il existe, pas que vous refassiez toute la dérivation mathématique.

Question 5

Quel est le rôle exact du IDMC dans cet essai ?

Réponse type

Le IDMC a pour rôle d'examiner de façon **indépendante** les résultats intermédiaires, en particulier les données d'efficacité et de sécurité, et de juger s'il est raisonnable de poursuivre, modifier ou arrêter l'essai. Son indépendance est essentielle car l'investigateur ou le promoteur ont des intérêts scientifiques, cliniques ou institutionnels qui ne doivent pas contaminer la décision. La bonne réponse insiste donc sur le fait que le IDMC protège à la fois les patients et l'intégrité scientifique de l'essai.

Question 6

Discutez l'intérêt d'une analyse intermédiaire de OS.

Réponse type

Une interim OS peut être utile parce que la survie globale reste le critère le plus "dur" et le plus convaincant cliniquement. Si un bénéfice net apparaît tôt, il peut être difficilement défendable de ne pas le prendre en compte. Mais l'inconvénient est majeur : les données d'OS sont souvent plus **immatures**, exposées à davantage de censure et à l'influence de traitements ultérieurs. La bonne copie conclut donc que l'interim OS est défendable, mais plus délicate à interpréter qu'une interim de PFS.

Question 7

Pourquoi la borne de OS est-elle plus stricte ?

Réponse type

La borne OS est plus stricte parce qu'on se situe dans une hiérarchie statistique où plusieurs temps de lecture et plusieurs objectifs d'efficacité coexistent. Une partie du risque alpha a déjà été allouée ailleurs, et l'OS est souvent plus tardive et plus sensible à l'immaturité des données. Ici, la borne effective $0,0124$ bilatérale est plus sévère que celle de l'interim PFS, ce qui traduit une protection renforcée du risque global de faux positif. En examen, il faut dire que **plus on regarde tôt ou plus la hiérarchie est complexe, plus la borne peut être exigeante**.

Question 8

Que peut-on conclure si une borne d'interim est franchie ?

Réponse type

Si une borne planifiée est franchie, on peut conclure qu'il existe un signal statistiquement compatible avec le niveau d'exigence fixé pour ce look précis. Cela ne signifie pas automatiquement qu'il faut arrêter l'essai dans tous les cas, mais cela rend une décision d'arrêt ou de communication beaucoup plus légitime. La bonne nuance est importante : **franchir une borne donne un droit méthodologique à conclure, mais la décision finale reste clinique, éthique et stratégique**.

Question 9

Pourquoi un résultat précoce peut-il surestimer la taille d'effet ?

Réponse type

Parce qu'un résultat observé tôt peut refléter en partie une fluctuation favorable, un nombre d'événements encore limité, ou une dynamique d'effet qui se stabilisera ensuite à un niveau moins spectaculaire. C'est un problème classique des arrêts précoces pour bénéfice. Une bonne réponse d'examen précise que cela n'invalide pas l'analyse intermédiaire, mais impose de discuter la **maturité des données** et la possibilité d'un **effet amplifié transitoirement**.

Question 10

Rédigez une réponse d'examen complète sur la stratégie d'analyses intermédiaires de cet essai.

Réponse type

L'essai comporte une véritable stratégie d'analyses intermédiaires, intégrée au protocole. Une première lecture était prévue sur le PFS après au moins 174 événements, avec contrôle du risque alpha global par une approche Lan-DeMets implémentant une frontière O'Brien-Fleming. Cela permettait de détecter tôt un bénéfice éventuel tout en évitant de multiplier les faux positifs liés aux regards successifs sur les données. Le papier donne un alpha effectif de 0,0337 bilatéral pour cette interim, puis une borne plus stricte pour l'interim d'OS (0,0124 bilatéral), ce qui est cohérent avec la hiérarchie des critères et la maturité plus délicate des données de survie globale. Les résultats intermédiaires étaient revus par un IDMC indépendant, garantissant une décision moins exposée aux biais du promoteur ou des investigateurs. La bonne conclusion est donc que la stratégie d'interims est méthodologiquement solide, mais que tout résultat précoce doit toujours être relu à la lumière de la maturité des données et du risque de surestimation de la taille d'effet.

6.2. 2026 EPI

Article et contexte

Article support : **Rosenström et al., Lancet Psychiatry 2025**, comparaison naturaliste finlandaise entre **guided iCBT** et **face-to-face CBT** pour la dépression.

- étude observationnelle rétrospective sur registres ;
- ATE estimé par TMLE avec Super Learner ;
- outcome principal : changement de PHQ-9 ;
- pas d'essai randomisé, pas de plan de looks successifs, pas de vraie logique confirmatoire d'analyses intermédiaires ;
- article très utile pour apprendre à dire **pourquoi** le cadre des interims n'est pas automatiquement transposable.

Questions fictives proposées

1. Pourquoi cette étude ne se prête-t-elle pas à de vraies analyses intermédiaires confirmatoires ?
2. Quelle différence faites-vous entre analyse intermédiaire et simple monitoring descriptif en cours de projet ?
3. Pourquoi n'y a-t-il pas ici d'alpha-spending au sens des essais ?
4. Quel rôle pourrait jouer un monitoring méthodologique dans une étude comme celle-ci ?
5. Pourquoi un DSMB n'aurait-il pas ici la même fonction que dans un essai ?
6. Si cette étude avait été conçue comme essai pragmatique, quel critère intermédiaire aurait pu être suivi ?

7. Quelles règles d'arrêt hypothétiques aurait-on pu imaginer pour sécurité ou futilité ?
8. Comment aurait-on géré la multiplicité si PHQ-9 avait été analysé à plusieurs looks ?
9. Pourquoi des analyses intermédiaires répétées seraient-elles délicates avec ATE / TMLE ?
10. Rédigez une réponse d'examen opposant essai avec interims et étude observationnelle sans vraie logique d'interim confirmatoire.

Extrait de l'article à analyser

Ce que l'examineur veut vérifier

- savez-vous dire clairement qu'une méthode n'est **pas transposable telle quelle** d'un design à l'autre ;
- savez-vous opposer essai randomisé confirmatoire et étude observationnelle causaliste ;
- savez-vous néanmoins construire un raisonnement hypothétique propre si l'on vous force à imaginer un essai pragmatique ;
- savez-vous rester ancré dans le papier réel sans inventer une méthodologie absurde.

Corrections détaillées

Question 1

Pourquoi cette étude ne se prête-t-elle pas à de vraies analyses intermédiaires confirmatoires ?

Réponse type

Parce qu'il ne s'agit pas d'un essai prospectif randomisé avec calendrier de looks successifs et contrôle séquentiel du risque alpha. Les auteurs exploitent ici des données de registre déjà constituées pour estimer un ATE par TMLE. On n'est donc pas dans une logique "on accumule les données puis on décide éventuellement d'arrêter tôt". La bonne réponse est : **ce design appelle surtout une discussion de validité causale, pas un plan d'analyses intermédiaires confirmatoires.**

Question 2

Quelle différence faites-vous entre analyse intermédiaire et simple monitoring descriptif en cours de projet ?

Réponse type

Une analyse intermédiaire implique une **règle de décision formelle** liée à l'efficacité, la futilité ou la sécurité, avec conséquences possibles sur la poursuite de l'étude. Un monitoring descriptif, lui, vérifie plutôt la qualité des données, le taux de données manquantes, la cohérence des flux, ou des indicateurs opérationnels. Dans cette étude, un monitoring descriptif aurait du sens ; une analyse intermédiaire confirmatoire, beaucoup moins. C'est précisément cette distinction que l'examineur peut vouloir faire émerger.

Question 3

Pourquoi n'y a-t-il pas ici d'alpha-spending au sens des essais ?

Réponse type

Parce qu'il n'existe pas ici de plan séquentiel de tests répétés sur un critère principal randomisé. L' α -spending sert à répartir le risque de première espèce lorsqu'on examine plusieurs fois des résultats d'essai avant la fin. Dans cet article, le cœur du travail est l'estimation d'un effet causal moyen sur des données observationnelles, avec analyses de sensibilité et machine learning. Le problème central n'est donc pas la dépense α entre looks successifs, mais la **confusion résiduelle**, les **hypothèses causales** et la **robustesse du modèle**.

Question 4

Quel rôle pourrait jouer un monitoring méthodologique dans une étude comme celle-ci ?

Réponse type

Un monitoring méthodologique pourrait suivre la qualité des chaînages de registres, la complétude du PHQ-9, la stabilité des définitions, la distribution des covariables, ou encore le comportement des modèles prédictifs. Autrement dit, on peut surveiller la **qualité de production de l'étude** sans pour autant parler d'analyses intermédiaires au sens confirmatoire. Cette nuance est importante : **tout projet gagne à être monitoré, mais tous les projets n'appellent pas un cadre séquentiel d'arrêt.**

Question 5

Pourquoi un DSMB n'aurait-il pas ici la même fonction que dans un essai ?

Réponse type

Dans un essai thérapeutique, le DSMB protège des participants exposés prospectivement à une intervention dont le bénéfice ou la toxicité peuvent imposer un arrêt. Ici, les patients ne sont pas randomisés prospectivement dans le cadre de l'étude ; les auteurs exploitent des soins déjà réalisés. Le besoin d'un comité de sécurité indépendant n'est donc pas du même ordre. On peut imaginer une gouvernance scientifique ou éthique des données, mais ce n'est pas la même mission qu'un DSMB d'essai.

Question 6

Si cette étude avait été conçue comme essai pragmatique, quel critère intermédiaire aurait pu être suivi ?

Réponse type

Dans un essai pragmatique comparant iCBT et fCBT, un critère intermédiaire naturel aurait pu être le **changement précoce de PHQ-9**, par exemple après quelques semaines ou quelques séances. On aurait aussi pu suivre un indicateur d'engagement thérapeutique ou d'abandon, s'il avait une portée clinique réelle. Le point important est de rester cohérent avec l'article réel : le PHQ-9 est déjà l'outcome central, donc c'est la meilleure base pour imaginer un interim fictif.

Question 7

Quelles règles d'arrêt hypothétiques aurait-on pu imaginer pour sécurité ou futilité ?

Réponse type

Dans un essai pragmatique fictif, on aurait pu prévoir des règles d'arrêt pour **futilité** si la probabilité d'observer une différence cliniquement pertinente entre iCBT et fCBT devenait très faible, ou pour **sécurité** si l'une des modalités s'accompagnait d'une aggravation nette, d'un risque suicidaire accru, ou

d'un abandon massivement supérieur. Ce raisonnement reste hypothétique, car l'article réel n'est pas conçu ainsi. Ce que l'examineur veut entendre, c'est que vous savez proposer des règles plausibles **sans prétendre qu'elles existaient réellement dans le papier.**

Question 8

Comment aurait-on géré la multiplicité si PHQ-9 avait été analysé à plusieurs looks ?

Réponse type

Si PHQ-9 avait été testé à plusieurs looks dans un essai, il aurait fallu prévoir une stratégie de contrôle du risque alpha : par exemple un alpha-spending de type Pocock, O'Brien-Fleming ou Lan-DeMets. Le principe aurait été le même que dans un essai de survie : **ne pas réutiliser naïvement le seuil usuel à chaque analyse.** Même si le critère est continu et non un temps jusqu'à événement, le problème statistique de multiplicité séquentielle reste réel.

Question 9

Pourquoi des analyses intermédiaires répétées seraient-elles délicates avec ATE / TMLE ?

Réponse type

Parce que TMLE est déjà une procédure causaliste sophistiquée reposant sur des modèles d'assignation et d'outcome, souvent alimentés par du machine learning. Répéter ce type d'estimation à plusieurs looks en temps réel soulèverait des questions complexes de stabilité, de reparamétrisation, de variabilité et de contrôle d'erreur. Le problème n'est donc pas seulement "combien de fois on regarde", mais aussi **comment l'estimateur lui-même se comporte dans un cadre séquentiel.** C'est une bonne question d'examen justement parce qu'elle force à distinguer la logique des essais classiques et celle des estimateurs causaux modernes.

Question 10

Rédigez une réponse d'examen opposant essai avec interims et étude observationnelle sans vraie logique d'interim confirmatoire.

Réponse type

Dans un essai randomisé confirmatoire, une analyse intermédiaire est une analyse planifiée avant la fin de l'étude, avec des règles explicites d'efficacité, de futilité ou de sécurité, et une stratégie de contrôle du risque alpha global. Elle peut conduire à poursuivre, modifier ou arrêter l'essai. Dans l'article EPI 2026, on n'est pas du tout dans ce cadre : les auteurs exploitent rétrospectivement des données de registre pour estimer un ATE par TMLE, sans calendrier de looks successifs ni logique d'arrêt précoce. Le bon raisonnement d'examen consiste donc d'abord à dire que le vocabulaire d'"analyse intermédiaire" est ici impropre au sens strict. On peut en revanche imaginer, dans un essai pragmatique fictif inspiré de cette comparaison, un suivi précoce de PHQ-9, des règles de sécurité ou de futilité et un contrôle de multiplicité. Mais cela relèverait d'un protocole différent de celui du papier réel.

7. Réponses prêtes à l'emploi

7.1. Si l'on vous demande de définir une analyse intermédiaire

Une analyse intermédiaire est une analyse planifiée avant l'analyse finale, prévue dans le protocole ou le plan d'analyse statistique, avec des règles explicites de décision et un contrôle

du risque alpha global.

7.2. Si l'on vous demande pourquoi il faut contrôler l'alpha

Regarder plusieurs fois les données sans correction augmente le risque de faux positif. C'est pourquoi un plan séquentiel doit utiliser une stratégie de type alpha-spending, afin de préserver le risque global de première espèce.

7.3. Si l'on vous demande le rôle du DSMB / IDMC

Le comité indépendant de monitoring examine les données intermédiaires, notamment d'efficacité et de sécurité, et aide à décider s'il faut poursuivre, modifier ou arrêter l'essai, sans dépendre directement des intérêts du promoteur ou des investigateurs.

7.4. Si l'on vous demande pourquoi ce raisonnement ne s'applique pas à une cohorte naturaliste

Dans une étude observationnelle rétrospective, il n'y a pas de randomisation prospective ni de calendrier de looks confirmatoires à protéger. Le problème principal n'est donc pas l'alpha-spending, mais la validité causale, la confusion et la qualité des données.

8. Checklist finale

Checklist de 20 secondes

- Ai-je bien dit s'il s'agit d'une vraie analyse intermédiaire ou non ?
- Ai-je distingué efficacité, futilité et sécurité ?
- Ai-je nommé la méthode de contrôle du risque alpha quand elle existe ?
- Ai-je dit pourquoi un seuil effectif d'interim n'est pas un simple seuil usuel répété ?
- Ai-je mentionné le rôle du IDMC / DSMB si nécessaire ?
- Ai-je terminé par une conclusion clinique et méthodologique ?

9. Conclusion

Pour traiter correctement une question d'examen sur les analyses intermédiaires, il faut retenir trois idées.

1. **Une analyse intermédiaire est un dispositif planifié, pas une curiosité statistique improvisée.**
2. **Sa validité dépend du contrôle du risque alpha et d'une gouvernance indépendante.**
3. **Ce raisonnement a du sens dans un essai prospectif; il n'est pas transposable tel quel à une étude observationnelle naturaliste.**

Le contraste entre RC 2026 et EPI 2026 est précisément utile pour l'examen : le premier article montre comment fonctionne une vraie stratégie d'interims, le second montre pourquoi il faut parfois refuser le mauvais cadre méthodologique avant même de discuter des détails techniques.