



Société Française de
Pharmacologie et de Thérapeutique

Groupe de Travail Méthodologie

Livre blanc SFPT

De la nécessité de la méthodologie
dans l'évaluation des médicaments

Document compagnon

Dossier 23 – Les estimands

11 mai 2022

Groupe de rédaction / relecture (Par ordre alphabétique)

- Theodora Angoulvant
- Laurent Bertolotti
- Jean-Luc Cracowski
- Michel Cucherat
- Dominique Deplanque
- Guillaume Grenet
- François Gueyffier
- Behrouz Kassai
- Charles Khouri
- Silvy Laporte
- Bruno Laviolle
- Jean-Christophe Lega
- Clara Locher
- Florian Naudet
- Edouard Ollier
- Antoine Pariente
- Matthieu Roustit
- Tabassome Simon



[Licence Creative Commons](#)

Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International

Vous êtes autorisé à :

- Partager — copier, distribuer et communiquer le matériel par tous moyens et sous tous formats
- Adapter — remixer, transformer et créer à partir du matériel pour toute utilisation, y compris commerciale.

Table des matières

1	Introduction.....	1
2	La quantité d'effet, une notion plurielle	2
3	Disposition réglementaire dans l'essai clinique	4
4	Les événements intercurrents.....	6
5	Estimands standards	8
5.1	Treatment Policy	8
5.2	Composite Strategy	9
5.3	Hypothetical Strategy (Trial Product Estimand).....	10
5.4	Principal Stratum Strategy.....	11
5.5	While on Treatment Strategy	12
5.6	Synthèse	Erreur ! Signet non défini.
6	REFERENCES	Erreur ! Signet non défini.

1 Introduction

Le concept d'estimand est à l'origine un concept statistique. Il désigne la quantité que l'on cherche à estimer. Il permet de distinguer l'objet de l'estimation (l'estimand) de l'outil utilisé (l'estimateur) et du résultat particulier obtenu sur un échantillon, l'estimation [1]. Un estimateur permet d'obtenir une estimation numérique de l'estimand.

L'essai thérapeutique cherche à estimer la taille d'effet qu'apporte un traitement. Cette quantité d'effet est l'objet de l'estimand. C'est la mesure que l'on aspire à faire avec l'essai entrepris. En fonction de la nature du critère de jugement et du plan d'analyse statistique différents estimateurs sont utilisables : calcul du risque relatif brut, différence des moyennes finales, différences des moindres carrés, hazard ratio produit par un modèle de Cox par exemple. Et l'essai aura comme résultat l'estimation numérique de cette quantité.

Toutefois, cette quantité d'effet, observée entre les résumés statistique du devenir des patients (comme la moyenne ou le taux d'évènements) de différents groupes de l'essai, est sujette à des variations directement liées à la stratégie d'analyse d'évènements intercurrents. L'estimand est donc un concept populationnel qui implique une estimation statistique. Comme il s'agit d'effet causal du traitement, l'estimand doit aussi correspondre à une question causale (cf. dossier sur l'inférence causale).

L'intérêt de cette conceptualisation est de faire apparaître qu'il existe possiblement plusieurs façons de concevoir la quantité d'effet, c'est-à-dire plusieurs estimands possibles, correspondant à différents points de vue ou façon de gérer les évènements intercurrents survenant lors d'un essai (cf. section suivante). La « quantité d'effet » n'est pas un concept trivial et nécessite, même si cela n'était pas vraiment fait jusqu'à présent, qu'elle soit soigneusement explicitée dans un essai (au niveau du protocole ou, éventuellement, du SAP).

2 La quantité d'effet, une notion plurielle

Imaginons l'essai thérapeutique d'un nouvel antihypertenseur dont l'objectif est d'estimer l'effet à 1 an du traitement sur la pression artérielle systolique (PAS) par rapport au placebo. L'effectif est de 200 patients répartis entre les 2 groupes. La différence entre les 2 groupes de la moyenne des PAS mesurées à 1 an est de -12 mmHg en faveur du nouveau traitement dans l'analyse en intention de traiter. Comme 10% des patients dans les deux groupes sont perdus de vue, un remplacement des données manquantes conservateur (BOCF) a été effectué dans cette analyse en ITT, pour protéger le résultat contre le biais d'attrition. L'effet obtenu par cette analyse (-12) n'est donc pas une mesure directe de l'effet de l'antihypertenseur, mais une mesure « transformée » pour des nécessités de contrôle de biais, et inférieure, par construction, à son réel effet. Sans ces perdus de vue, l'effet obtenu aurait été un supérieur, peut-être -14 mm Hg. Déjà apparaissent deux quantités d'effet : la vraie quantité apportée par le traitement (telle qu'elle serait mesurée dans un essai idéal sans aucun perdu de vue ou autres événements intercurrents, cf. infra) et celle apportée par une analyse en ITT. L'analyse per protocol définit une autre quantité d'effet (un autre estimand donc).

Malencontreusement, du fait d'une erreur de conditionnement, 40 patients du groupe expérimental ont reçu en fait du placebo ! La quantité d'effet mesurée en ITT, analysée en fonction du traitement alloué et non pas du traitement reçu, est de ce fait très inférieure au vrai effet du traitement.

Cette erreur de conditionnement a toutes les chances d'être entièrement aléatoire. Comme il n'y a aucune raison qu'elle survienne en fonction de la PAS des patients concernés, par exemple préférentiellement chez les patients ayant des valeurs très élevées, la non-prise en compte de ces sujets n'est pas susceptible d'entraîner un biais (au sens qualitatif, faisant conclure à tort à l'effet non nul du traitement). En revanche leur maintien dans l'analyse entraîne la mesure d'une quantité qui est l'effet du traitement dans un essai dont 40% des patients sous traitement ont reçu un placebo ! Cette estimand n'a pas de sens médical, car cette erreur n'a aucune chance de se reproduire dans le futur, en pratique médicale. Il n'est donc peut-être pas nécessaire d'intégrer, dans la mesure de l'apport du produit, les conséquences d'événements intercurrents qui ne se reproduiront jamais. Cette situation s'oppose, par exemple, à celle des effets indésirables qui entraînent l'arrêt du traitement. Le maintien de ces patients dans l'analyse en ITT conduit lui aussi à une minoration de la quantité d'effet mesurée, mais ces sujets doivent impérativement être pris en considération, car il s'agit d'une conséquence directe du traitement (une de ses propriétés). Ces effets indésirables se reproduiront en pratique chez les patients traités et il est donc indispensable d'intégrer leur conséquence dans l'appréciation de l'effet du produit, car elles sont susceptibles de réduire cet intérêt.

Au passage, notons qu'il serait très fréquent de dire que, dans cet exemple des erreurs de conditionnement, le résultat ITT est biaisé, car différent quantitativement de celui attendu dans un essai sans cette erreur de conditionnement. En fait ce n'est pas un biais, car le résultat est conforme à ce qu'il devrait être dans cette situation et avec l'estimand en ITT pure (l'essai est considéré comme étant parfaitement à l'abri des biais et mesurant sans erreur autre qu'aléatoire, ce qu'il doit mesurer). Cet exemple montre bien l'intérêt de conceptualiser la notion d'estimand. Un résultat peut ne pas être approprié à la décision, car il est faux en raison d'un biais, mais aussi parce qu'il n'appréhende pas correctement la quantité d'effet nécessaire à cette prise de décision (comme dans une analyse excluant les arrêts de traitement pour effet indésirable).

L'estimand peut être vu comme une étape conceptuelle entre l'objectif de l'essai et le design de l'étude afin que celui-ci, comme l'analyse statistique, réponde au mieux à la question de l'étude [2].

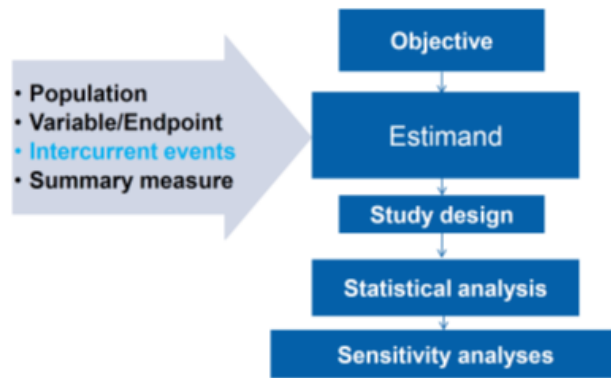


Figure 1 – Place de l'estimand dans le processus de réalisation d'un essai

3 Disposition réglementaire dans l'essai clinique

L'estimand a récemment fait son apparition au niveau réglementaire pour l'essai thérapeutique d'enregistrement des médicaments sous la forme d'une annexe à ICH E9¹ [3], effectivement depuis environ 2020. En conséquence, les premiers essais ayant mis en œuvre ce concept commencent juste à être publiés [4, 5, 6, 7, 8, 9, 10].

Exemple de mentions aux estimands rencontrées dans quelques publications récentes

The co-primary endpoints were investigator-assessed progression-free survival and overall survival, analysed according to the treatment policy estimand (an equivalent of the intention-to-treat principle) [4].

Two efficacy-related estimands were defined: treatment policy (regardless of treatment discontinuation or rescue medication) and trial product (on treatment without rescue medication) in all participants randomly assigned [5].

Dans cet addendum, la formalisation de l'estimand va au-delà du concept purement statistique et recouvre plusieurs dimensions constitutives qui sont :

- Les traitements comparés,
- La population visée (la totalité de l'essai ou un sous-groupe),
- Le critère de jugement
- Le type statistique du résumé populationnel utilisé (population-level summary) : différence des moyennes, différences des changements, hazard ratio, etc.,
- Et surtout la gestion des événements intercurrents (cf. section 4).

De plus en plus de protocoles présentent la définition des estimands qu'ils utiliseront sous la forme d'un tableau reprenant ces différentes dimensions. En effet, chaque objectif ou sous-objectif d'un essai correspond à un estimand différent.

Le protocole de l'essai du vaccin COVID-19 Pfizer/BioNtech² présente ses estimands de la manière suivante :

¹ https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf

² "A Phase 1/2/3, Placebo-Controlled, Randomized, Observer-Blind, Dose-Finding Study to Evaluate the Safety, Tolerability, Immunogenicity, and Efficacy of SARS-COV-2 RNA Vaccine Candidates Against COVID-19 in Healthy Individuals":

3.2. For Phase 2/3

Objectives ^a	Estimands	Endpoints
Primary Efficacy		
To evaluate the efficacy of prophylactic BNT162b2 against confirmed COVID-19 occurring from 7 days after the second dose in participants without evidence of infection before vaccination	In participants complying with the key protocol criteria (evaluable participants) at least 7 days after receipt of the second dose of study intervention: $100 \times (1 - \text{IRR})$ [ratio of active vaccine to placebo]	COVID-19 incidence per 1000 person-years of follow-up based on central laboratory or locally confirmed NAAT in participants with no serological or virological evidence (up to 7 days after receipt of the second dose) of past SARS-CoV-2 infection
To evaluate the efficacy of prophylactic BNT162b2 against confirmed COVID-19 occurring from 7 days after the second dose in participants with and without evidence of infection before vaccination	In participants complying with the key protocol criteria (evaluable participants) at least 7 days after receipt of the second dose of study intervention: $100 \times (1 - \text{IRR})$ [ratio of active vaccine to placebo]	COVID-19 incidence per 1000 person-years of follow-up based on central laboratory or locally confirmed NAAT
Primary Safety		
To define the safety profile of prophylactic BNT162b2 in <u>the first 360 participants</u> randomized (Phase 2)	In participants receiving at least 1 dose of study intervention, the percentage of participants reporting: <ul style="list-style-type: none"> Local reactions for up to 7 days following each dose Systemic events for up to 7 days following each dose AEs from Dose 1 to 7 days after the second dose SAEs from Dose 1 to 7 days after the second dose 	<ul style="list-style-type: none"> Local reactions (pain at the injection site, redness, and swelling) Systemic events (fever, fatigue, headache, chills, vomiting, diarrhea, new or worsened muscle pain, and new or worsened joint pain) AEs SAEs
To define the safety profile of prophylactic BNT162b2 in <u>all participants</u> randomized in Phase 2/3	In participants receiving at least 1 dose of study intervention, the percentage of participants reporting: <ul style="list-style-type: none"> Local reactions for up to 7 days following each dose Systemic events for up to 7 days following each dose AEs from Dose 1 to 1 month after the second dose SAEs from Dose 1 to 6 months after the second dose 	<ul style="list-style-type: none"> AEs SAEs In a subset of at least 6000 participants: <ul style="list-style-type: none"> Local reactions (pain at the injection site, redness, and swelling) Systemic events (fever, fatigue, headache, chills, vomiting, diarrhea, new or worsened muscle pain, and new or worsened joint pain)
To define the safety profile of prophylactic BNT162b2 in participants 12 to 15 years of age in Phase 3	In participants receiving at least 1 dose of study intervention, the percentage of participants reporting: <ul style="list-style-type: none"> Local reactions for up to 7 days following each dose Systemic events for up to 7 days following each dose AEs from Dose 1 to 1 month after the second dose SAEs from Dose 1 to 6 months after the second dose 	<ul style="list-style-type: none"> Local reactions (pain at the injection site, redness, and swelling) Systemic events (fever, fatigue, headache, chills, vomiting, diarrhea, new or worsened muscle pain, and new or worsened joint pain) AEs SAEs

L'estimand permet ainsi de préciser, pour chacun des objectifs de l'essai, la population et la stratégie d'analyse du critère de jugement qui lui est associé. Alors que l'approche classique consistait à décrire plusieurs populations d'analyse (ITT, ITT modifiée, sécurité) permettant des analyses de sensibilité, l'estimand intègre la population d'analyse pertinente pour un objectif donné. On voit par exemple dans le tableau ci-dessus que les objectifs de sécurité sont exclusivement analysés chez les patients ayant reçu au moins une dose, ce qui correspond au véritable objectif.

4 Les évènements intercurrents

L'apport de l'approche basée sur les estimands provient principalement de l'obligation de préciser de manière explicite comment seront gérés les évènements intercurrents qui émaillent le suivi des patients dans tout essai thérapeutique. Ces évènements intercurrents interfèrent avec la mesure du critère de jugement et sont de différents types :

- arrêt du traitement de l'étude pour effet indésirable,
- arrêt du traitement pour recourir à un autre traitement (commercialisé),
- utilisation d'un traitement de secours ou d'un traitement concomitant,
- décès avant l'observation du critère de jugement de l'étude,
- décès sans rapport avec la maladie,
- écart au protocole, erreur,
- mauvaise observance,
- retrait de consentement,
- exclusion post-randomisation pour non-respect d'un critère d'inclusion exposant le patient à un sur-risque (ex : contre-indication au traitement)
- données manquantes post-randomisation

Ces évènements diffèrent légèrement suivant si l'essai est court ou long et si le critère de jugement est une mesure quantitative (variable continue) ou un évènement clinique (variable binaire).

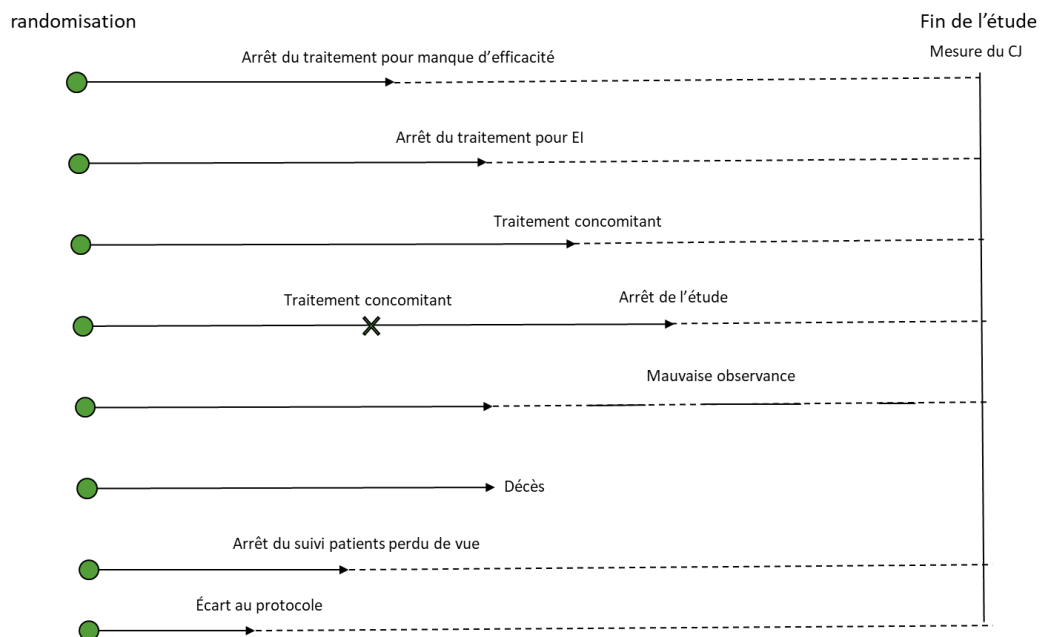


Figure 2 – Les différents évènements intercurrents qui peuvent émailler le suivi des patients dans tout essai thérapeutique.

Tous ces évènements ont en commun de poser des questions sur comment analyser les patients concernés : les conserver, les « sortir d'étude », imputer les valeurs manquantes, etc. Ces questions étaient déjà gérées dans les protocoles avant l'apparition du concept d'estimand mais parfois avec un manque de précision, ou une information disséminée à travers plusieurs sections du protocole. Le concept d'estimand implique de décrire de manière formelle et standardisée tous ces éléments (dans

une nouvelle section du protocole), et ce pour chaque objectif de l'essai. Ainsi, il apparaît plus clairement dans le protocole que chaque objectif fait l'objet d'une analyse avec des règles qui lui sont propres, afin de répondre plus précisément à la question posée

La notion d'estimand n'introduit rien de nouveau [11] dans la méthodologie des essais cliniques, mais permet simplement plus d'anticipation, de transparence et de standardisation.

5 Estimands standards

L'addendum au document ICH E9 définit plusieurs estimands « standards », sans exclure que d'autres puissent être utiles et utilisés.

Dans la suite de ce dossier, ces différents estimands vont être présentés à partir de la situation hypersimplifiée représentée dans la Figure 3.

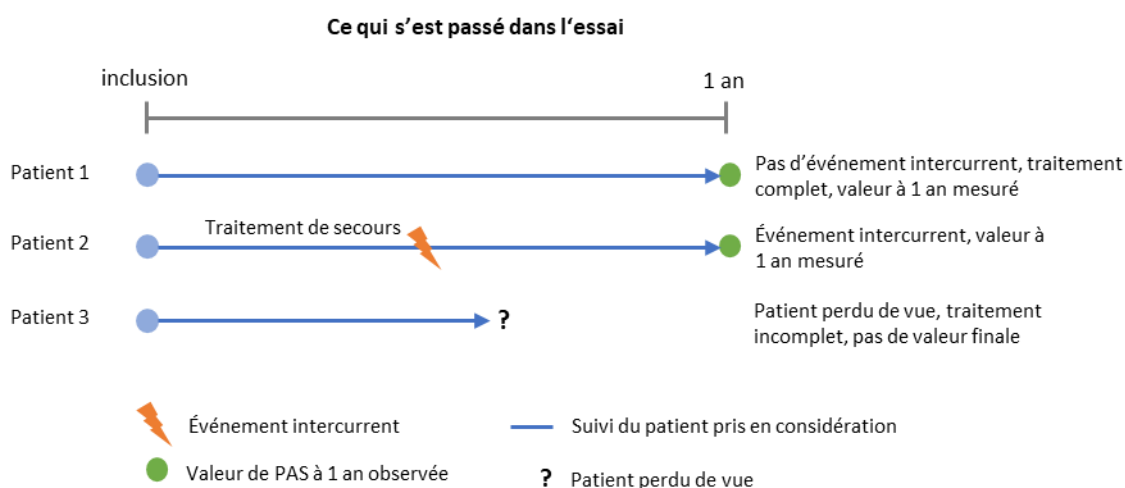


Figure 3 – Suivi de 3 patients représentant les différents cas de figure dont la prise en compte variera en fonction de l'estimand considéré (cf. les figures suivantes).

5.1 L'estimand « *Treatment Policy* »

Dans l'estimand « *treatment policy* », la survenue des événements intercurrents ne change rien dans la prise en compte des valeurs mesurées chez les patients. Ces événements sont considérés comme « normaux », survenant habituellement dans le traitement plus ou moins prolongé des patients et leurs conséquences doivent impérativement être intégrées à l'évolution du traitement. Il s'agit par exemple des arrêts pour effet indésirable.

Dans cet estimand, l'effet d'intérêt correspond à celui mesurable à la fin prévue de l'essai, quel que soit ce qui s'est produit durant le suivi (arrêt de traitement, traitement de secours, traitement concomitant, etc.). Pour les patients perdus de vue, la donnée manquante du critère de jugement à la fin du suivi programmé est imputée et remplacée, selon une approche conservatrice (en supériorité).

Cet estimand mesure en quoi la recommandation d'utiliser de ce traitement pourrait changer l'état de santé de la population de malades concernés. Il correspond à l'objectif des essais cliniques pivots qui est de déterminer l'utilité médicale d'une stratégie thérapeutique intégrant le nouveau traitement (et non pas de déterminer l'effet intrinsèque du traitement, qui correspond à l'estimand « *hypothetical* »).

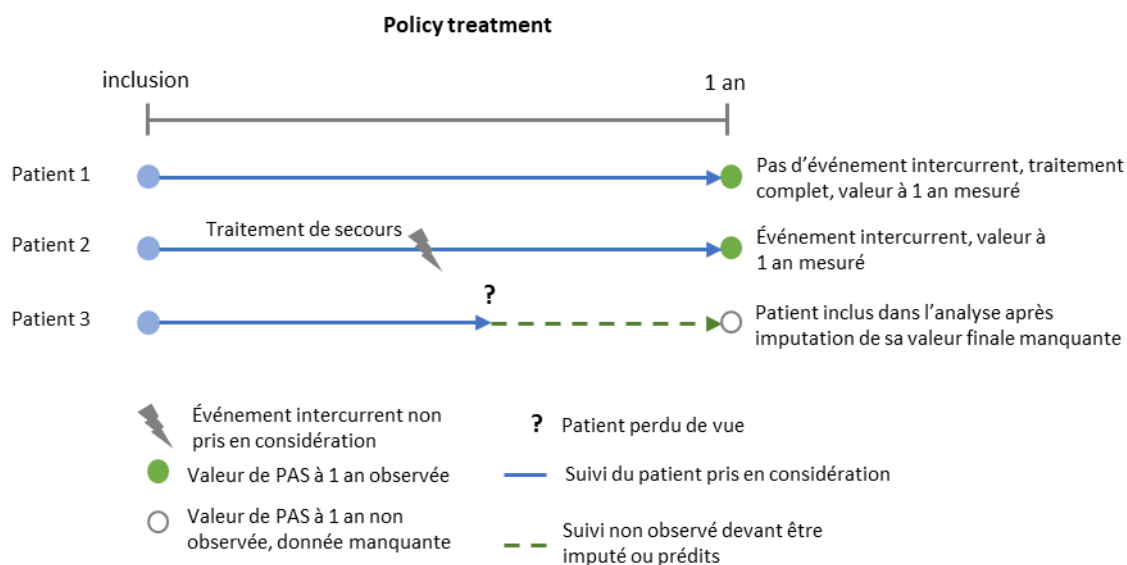


Figure 4 – Illustration de l'estimand « *treatment policy* ».

5.2 L'estimand « *Composite Strategy* »

Dans l'estimand « *composite strategy* », l'évènement intercurrent est pris en considération dans le critère de jugement qui est un composite du critère de jugement cible et de la survenue d'un évènement intercurrent. C'est par exemple un critère d'échec thérapeutique qui regroupe la non-obtention de l'objectif thérapeutique (un seuil d'EVA avec un antalgique par exemple) ou la nécessité de recourir à un traitement de secours (recours à l'antalgique de secours).

Proportion of patients with improvement in nasal polyps score of ≥ 1 point and completion of treatment period without surgery [10] (cité par [12]).

the proportion of patients in both groups who reached an absolute HbA1c value of $< 7\%$ at the end of the study without the use of rescue medication (cité par [12]).

Il n'y a donc plus de nécessité de mesurer ou d'imputer la valeur du critère de jugement manquante. Le dénominateur reste inchangé et correspond à la taille initiale des groupes (comme en ITT, *full set analysis*).

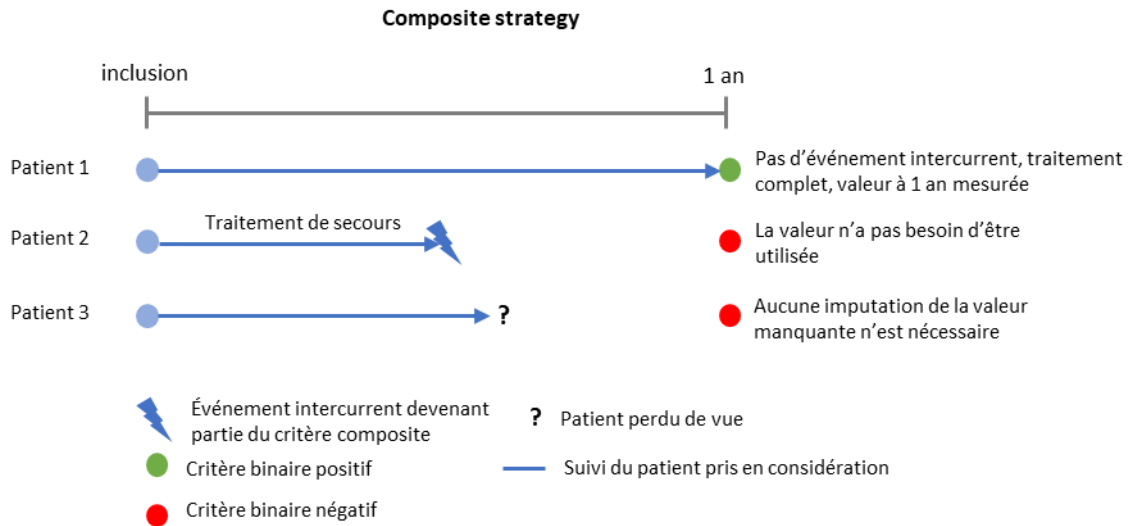


Figure 5 – Illustration de la prise en charge des différents cas de figure de l'exemple de la Figure 3 par l'estimand « composite strategy »

5.3 L'estimand « Hypothetical Strategy » (Trial Product Estimand)

L'estimand « hypothetical », aussi appelé « Trial Product Estimand », va chercher à mesurer l'efficacité intrinsèque du traitement. Aucune conséquence des événements intercurrents n'est prise en considération. Ces événements intercurrents sont considérés comme apportant simplement du bruit et sont donc nuisibles à la quantification de l'effet propre, théorique, du traitement.

Par exemple, si un patient switch en cours d'étude pour un traitement déjà commercialisé et cela sous la seule pression de la communication promotionnelle de ce traitement, ce switch va perturber l'estimation de l'effet propre du traitement étudié alors qu'il n'a aucune relation avec celui-ci. La valeur du critère de jugement qui aurait dû être obtenue à la fin du suivi, en l'absence de ce switch, va être prédite. Cette prédiction sera, par exemple, obtenue en extrapolant par un modèle l'évolution observée chez ce patient avant le switch, comme si, hypothétiquement, cet événement n'était pas survenu et si le patient avait été traité jusqu'à la fin de la période de suivi.

Si le patient décède de cause indépendante de la maladie, le même raisonnement est appliqué en considérant que ce décès est un événement aléatoire, complètement indépendant de l'objectif de l'essai (à la manière des risques compétitifs).

Cet estimand mesure donc l'effet hypothétique du traitement dans un essai où aucun événement intercurrent ne surviendrait. Il s'intéresse à une efficacité théorique intrinsèque du traitement, très différente du bénéfice clinique qu'est susceptible d'apporter le médicament (cf. dossier compagnon 2).

Cependant, si l'événement intercurrent est lié au traitement étudié, comme un effet indésirable, cet estimand va quantifier un l'effet hypothétique qui serait obtenu si le traitement n'entraînait aucun effet indésirable ou problème de tolérance. Dans ce cas, ce résultat n'a aucune réalité tangible et est purement théorique, car en pratique des événements indésirables surviendront.

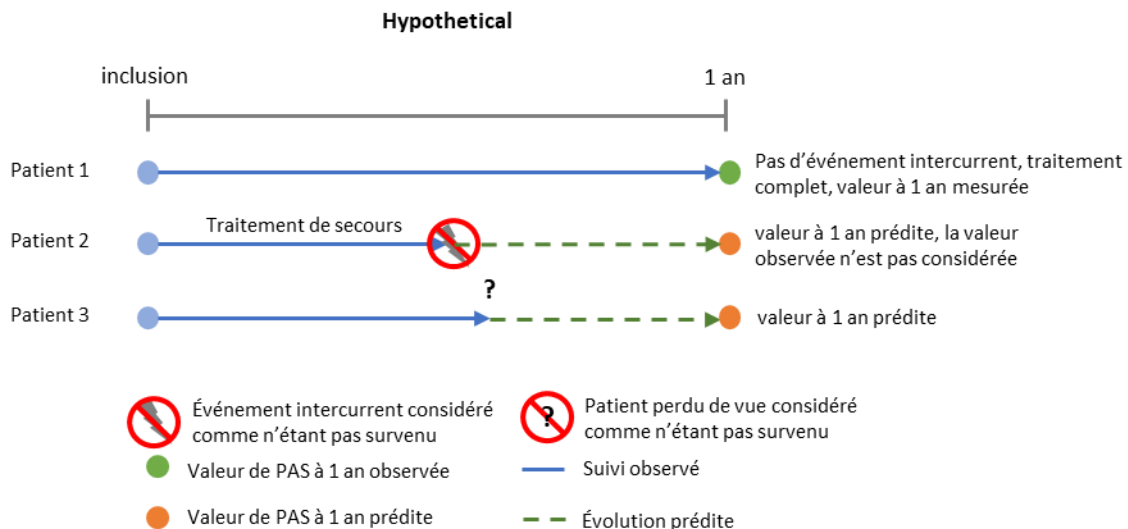
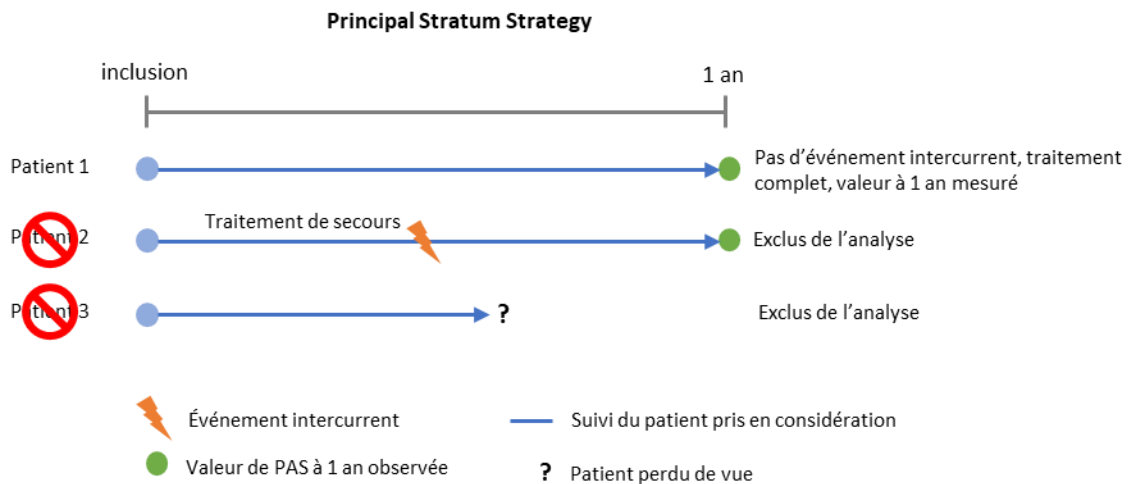


Figure 6 – Prise en charge des événements intercurrents et des patients perdus de vue par l'estimand « hypothetical ».

5.4 L'estimand "Principal Stratum Strategy"

L'estimand "Principal Stratum Strategy" ne considère que la sous population des patients de l'essai qui n'ont fait aucun événement intercurrent. Contrairement aux autres estimands, il est le seul à jouer sur les effectifs, les autres estimands retenant la totalité des patients randomisés. Il s'apparente de ce fait à l'analyse per protocol.

Dans d'autres cas, la strate retenue peut être celle des patients ayant présenté un certain événement intercurrent, par exemple, pour mesurer le bénéfice du traitement chez les répondeurs (définis d'après un critère intermédiaire) ou chez les sujets tolérant le traitement. Dans ce cas, comme la population d'analyse va être sélectionnée d'après des événements survenant après randomisation, un biais de sélection (au sens des études observationnelles et non par du ROB 1.0) est possible (par restriction sur un collider).



5.5 L'estimand "While on Treatment Strategy"

L'estimand "While on Treatment Strategy" va utiliser, pour les patients ayant arrêté le traitement de l'étude, la dernière valeur du critère de jugement mesurée sous traitement (et non pas la valeur réelle obtenue à la fin du suivi, comme avec l'estimand « policy treatment », ou la valeur extrapolée avec l'estimand « hypothetical »). Il quantifie donc une autre sorte d'efficacité théorique, intrinsèque du traitement, similaire à l'approche en per protocol.

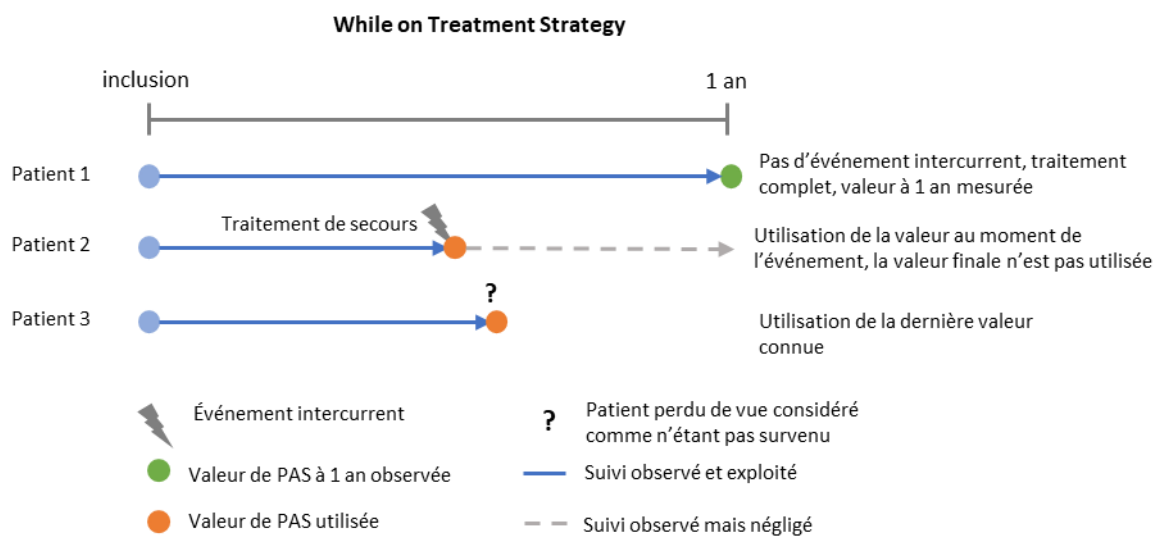


Figure 7 – Gestion des évènements intercurrents et des perdues de vue par l'estimand « on treatment »

Cet estimand peut être très trompeur sur le réel bénéfice que les patients pourraient attendre de l'initiation du traitement. Par exemple avec un produit entraînant une perte de poids, mais qui est très

mal toléré, sous traitement les patients perdent rapidement du poids, mais ils sont aussi obligés de l'arrêter dans 50% des cas, en moyenne au bout de 6 mois. L'estimand « *while on treatment* » peut donner une valeur flatteuse de perte de poids, mais cet estimand ne prend pas en considération qu'après l'arrêt du traitement, les patients reprennent leur poids d'origine, et ne peuvent ainsi tirer aucun bénéfice de leur incursion thérapeutique. En terme clinique, ce qui compte c'est une perte de poids qui se maintient dans le temps (pour espérer un bénéfice CV, locomoteur, etc.). L'estimand d'intérêt est donc un estimand qui s'intéresse à la perte de poids au long court, au moins 1 an par exemple. Celui-ci donnera alors une vision moins flatteuse du bénéfice pondéral de ce traitement, en intégrant les valeurs revenues au poids initial pour les 50% de patients ayant arrêté pour intolérance. Cet estimand approprié mettra en évidence le faible intérêt clinique que présente ce produit, ce que ne fera pas l'estimand « *while on treatment* ».

6 Lecture critique

L'intérêt de ces différents estimands n'est pas identique et varie en fonction de l'interlocuteur considéré [2].

Pour la régulation type AMM, la question posée est « quel est l'effet du traitement durant le temps où les patients le prennent ? ». À ce niveau les estimand « *while on treatment* » et « *hypothetical* » sont adaptés comme aussi le « *treatment policy* » si la question concerne plus l'utilité médicale.

Pour le patient lui-même, la question est « à quel bénéfice puis-je m'attendre si je débute ce traitement ? ». Seul l'estimand « *treatment policy* » apporte la réponse à cette question.

Pour les payeurs et les régulateurs à l'accès aux patients des médicaments, la question est « quelle amélioration de l'état de santé de la population des patients sera obtenue si l'on recommande l'utilisation de ce traitement dans la stratégie thérapeutique ? ». La réponse est aussi apportée par l'estimand « *treatment policy* ».

Pour le chercheur, la question est plutôt « quelle est l'action intrinsèque de cette molécule ? » et l'estimand « *hypothetical* » est alors adapté

Au total, il apparaît que pour la pratique médicale le seul estimand d'intérêt est « *treatment policy* » qui correspond d'ailleurs en tout point à l'analyse en intention de traiter avec remplacement conservateur des données manquantes qui était jusqu'à présent utilisé.

7 Références

8 REFERENCES

- 1 Little RJ, Lewis RJ. Estimands, Estimators, and Estimates. *JAMA* 2021 doi:10.1001/jama.2021.2886; PMID:34424281;
- 2 Phillips A, Clark T. Estimands in practice: Bridging the gap between study objectives and statistical analysis. *Pharm Stat* 2021;20:68–76 doi:10.1002/pst.2056;
- 3 INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE. ADDENDUM ON ESTIMANDS AND SENSITIVITY ANALYSIS IN CLINICAL TRIALS TO THE GUIDELINE ON STATISTICAL PRINCIPLES FOR CLINICAL TRIALS E9(R1).
- 4 Fennell DA, Ewings S, Ottensmeier C, et al. Nivolumab versus placebo in patients with relapsed malignant mesothelioma (CONFIRM): a multicentre, double-blind, randomised, phase 3 trial. *The Lancet Oncology* 2021;22:1530–40 doi:10.1016/S1470-2045(21)00471-X;
- 5 Mosenzon O, Blicher TM, Rosenlund S, et al. Efficacy and safety of oral semaglutide in patients with type 2 diabetes and moderate renal impairment (PIONEER 5): a placebo-controlled, randomised, phase 3a trial. *Lancet Diabetes Endocrinol* 2019;7:515–27 doi:10.1016/S2213-8587(19)30192-5; PMID:31189517;
- 6 Pieber TR, Bode B, Mertens A, et al. Efficacy and safety of oral semaglutide with flexible dose adjustment versus sitagliptin in type 2 diabetes (PIONEER 7): a multicentre, open-label, randomised, phase 3a trial. *Lancet Diabetes Endocrinol* 2019;7:528–39 doi:10.1016/S2213-8587(19)30194-9; PMID:31189520;
- 7 Lingvay I, Catarig A-M, Frias JP, et al. Efficacy and safety of once-weekly semaglutide versus daily canagliflozin as add-on to metformin in patients with type 2 diabetes (SUSTAIN 8): a double-blind, phase 3b, randomised controlled trial. *Lancet Diabetes Endocrinol* 2019;7:834–44 doi:10.1016/S2213-8587(19)30311-0;
- 8 Rabe KF, Martinez FJ, Ferguson GT, et al. Triple Inhaled Therapy at Two Glucocorticoid Doses in Moderate-to-Very-Severe COPD. *New Engl J Med* 2020;383:35–48 doi:10.1056/nejmoa1916046;
- 9 Rosenstock J, Bajaj HS, Janež A, et al. Once-Weekly Insulin for Type 2 Diabetes without Previous Insulin Treatment. *New Engl J Med* 2020;383:2107–16 doi:10.1056/nejmoa2022474;
- 10 Heerspink HJL, Parving H-H, Andress DL, et al. Atrasentan and renal events in patients with type 2 diabetes and chronic kidney disease (SONAR): a double-blind, randomised, placebo-controlled trial. *The Lancet* 2019;393:1937–47 doi:10.1016/S0140-6736(19)30772-X; PMID:30995972;
- 11 Rosenkranz G. Estimands-new statistical principle or the emperor's new clothes? *Pharm Stat* 2017;16:4–5 doi:10.1002/pst.1792; PMID:27966259;
- 12 Gogtay NJ, Ranganathan P, Aggarwal R. Understanding estimands. *Perspect Clin Res* 2021;12:106–12 doi:10.4103/picr.picr_384_20; PMID:34012908;