



Comité de surveillance indépendant dans les essais cliniques : de la justification scientifique à l'organisation

Clara Locher, Silvy Laporte, Peggy Derambure, Olivier Chassany, Cécile Girault,
Alix Avakiantz, Claire Bahans, Dominique Deplanque, Pierre Fustier, Anne
Françoise Germe, et al.

► To cite this version:

Clara Locher, Silvy Laporte, Peggy Derambure, Olivier Chassany, Cécile Girault, et al.. Comité de surveillance indépendant dans les essais cliniques : de la justification scientifique à l'organisation. *Thérapies*, 2024, 79 (1), pp.99-110. <10.1016/j.therap.2023.10.013>. <hal-04286449>

HAL Id: hal-04286449

<https://hal.science/hal-04286449v1>

Submitted on 8 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

Journal Pre-proof

Comité de surveillance indépendant dans les essais cliniques : de la justification scientifique à l'organisation

Clara Locher Silvy Laporte Peggy Derambure Olivier Chassany
Cécile Girault Alix Avakiantz Claire Bahans Dominique Deplanque
Pierre Fustier Anne-Françoise Germe Behrouz Kassaï Louis
Lacoste Nadine Petitpain Matthieu Roustit Tabassome Simon
Cécile Train Michel Cucherat



PII: S0040-5957(23)00179-8

DOI: <https://doi.org/doi:10.1016/j.therap.2023.10.013>

Reference: THERAP 834

To appear in: *Therapies*

Received Date: 19 October 2023

Accepted Date: 27 October 2023

Please cite this article as: Locher C, Laporte S, Derambure P, Chassany O, Girault C, Avakiantz A, Bahans C, Deplanque D, Fustier P, Germe A-Françoise, Kassaï B, Lacoste L, Petitpain N, Roustit M, Simon T, Train C, Cucherat M, Comité de surveillance indépendant dans les essais cliniques : de la justification scientifique à l'organisation, *Therapies* (2023), doi: <https://doi.org/10.1016/j.therap.2023.10.013>

THERAPIES

Ateliers de Giens 2023/Aspects organisationnels et réglementaires

Comité de surveillance indépendant dans les essais cliniques : de la justification scientifique à l'organisation*

Clara Locher^{a,*}, Silvy Laporte^b, Peggy Derambure^c, Olivier Chassany^d, Cécile Girault^e, Alix Avakiantz^{f,1}, Claire Bahans^{g,1}, Dominique Deplanque^{h,1}, Pierre Fustier^{i,1}, Anne-Françoise Germe^{j,1}, Behrouz Kassai^{k,1}, Louis Lacoste^{l,1}, Nadine Petitpain^{m,1}, Matthieu Roustit^{n,1}, Tabassome Simon^{o,1}, Cécile Train^{p,1}, Michel Cucherat^{q,1}

^a CHU Rennes, Inserm, Centre d'investigation clinique de Rennes (CIC1414), service de pharmacologie clinique, Institut de recherche en santé, environnement et travail (Irset), UMR S 1085, 35000 Rennes, France

^b Univ. Jean Monnet Saint-Etienne, UMR 1059 Inserm, URC/Pharmacologie Clinique, CHU Saint-Etienne, 42055 Saint-Etienne, France

^c AbbVie, 94528 Rungis, France

^d Unité de recherche clinique en économie de la santé (URC-ECO), hôpital Hôtel-Dieu, AP-HP, 75004 Paris, France

^e Fédération francophone de cancérologie digestive (FFCD), 21000 Dijon, France

^f Novartis Pharma S.A.S., 92500 Rueil-Malmaison, France

^g CHU de Limoges, Département de pédiatrie, 87000 Limoges, France

^h Univ. Lille, Inserm, CHU Lille, CIC 1403, Centre d'investigation clinique, 59000 Lille, France

ⁱ Département de recherche et développement clinique - Hématologie - BeiGene, Switzerland GmbH, 4051, Basel, Switzerland

^j CHU Lille, Institut de pharmacie, 59000 Lille, France

^k Centre d'investigation clinique 1407 INSERM - Hospices civils de Lyon, UMR 5558 CNRS université de Lyon, service de pharmacotoxicologie, 69000 Lyon, France

^l Pôle USSAR, anesthésie réanimation, CHU de Poitiers, 86021 Poitiers, France

^m *Unité de vigilance des essais cliniques, DRCI CHRU de Nancy, 54500 Vandoeuvre lès Nancy, France*

ⁿ *Univ. Grenoble Alpes, Inserm, CHU de Grenoble, CIC1406, 38000 Grenoble, France*

^o *Sorbonne Université, service de pharmacologie clinique, plateforme de recherche de l'est parisien (URCEST-CRCEST-CRB), Assistance publique - Hôpitaux de Paris (AP-HP), Hôpital St Antoine, 75000 Paris, France*

^p *Clinsearch, 92240 Malakoff, France*

^q *metaEvidence.org, service de pharmacotoxicologie Hospices civils de Lyon, 69000 Lyon, France*

Reçu le 19 octobre 2023 ; accepté le 27 octobre 2023

***Auteur correspondant.** CHU Rennes, Inserm, Centre d'investigation clinique de Rennes (CIC1414), service de pharmacologie clinique, Institut de recherche en santé, environnement et travail (Irset), UMR S 1085, 35000 Rennes, France.

Adresse e-mail : clara.locher@chu-rennes.fr (C. Locher)

¹ Les participants à la table ronde « Aspects organisationnels et réglementaires » des Ateliers de Giens 2023.

*Les articles, analyses et propositions issus des Ateliers de Giens sont ceux des auteurs et ne préjugent pas des propositions de leur organisation

Résumé

Les essais cliniques durent souvent plusieurs mois voire plusieurs années. Au fur et à mesure de l'avancée de l'essai, il peut être tentant de s'assurer que les données accumulées ne permettent pas déjà de répondre à la question posée par l'essai et ainsi arrêter précocement les inclusions ou le suivi. Mais de façon contre-productive, la connaissance et la prise en compte de résultats intermédiaires peuvent, dans certaines conditions, compromettre l'intégrité des résultats. C'est pour limiter ce risque – et assurer ainsi une fiable évaluation des thérapeutiques – que cette surveillance

de critères de sécurité et/ou d'efficacité en cours d'étude est confiée à un comité de surveillance indépendant. A partir des résultats qui leur sont transmis de façon confidentielle, le comité de surveillance indépendant évalue la balance du bénéfice-risque du traitement à l'étude et établit une recommandation quant à la poursuite, la modification ou l'arrêt de l'étude. Au travers de ces recommandations, les membres des comités de surveillance indépendant ont une responsabilité importante : une décision d'arrêt trop hâtive peut rendre l'essai non concluant et infructueux pour répondre à la question initiale et au contraire, une décision d'arrêt trop tardive peut exposer les participants à des interventions potentiellement inefficaces voire nocives. La mission confiée aux membres des comités de surveillance indépendant est donc particulièrement complexe.

Dans ce contexte, la table ronde des ateliers de Giens a été l'occasion de revenir sur la justification scientifique vis-à-vis de l'organisation des comités indépendants de surveillance et de rappeler la nécessité pour les membres des comités de surveillance indépendant d'être parfaitement formés aux problématiques innervantes aux analyses multiples, à l'obligation de confidentialité vis-à-vis des résultats et au fait que les recommandations d'arrêt doivent reposer sur des données suffisamment convaincantes pour évaluer la balance bénéfice-risque du traitement étudié.

MOTS CLÉS

Comité de surveillance indépendant ; Balance bénéfice-risque ; Analyses intermédiaires ; Intégrité

Abréviations

ANSM : Agence nationale de sécurité des médicaments et des produits de santé

BPC : bonnes pratiques cliniques

CSI : Comité de surveillance indépendant

DMC : *Data Monitoring Committee*

DSMB : *Data Safety and Monitoring Board*

EMA : *European Medicines Agency*

FDA : *Food and Drug Administration*

HAS: Haute autorité de santé

NIH : *National Institutes of Health*

OS : *overall survival*

PFS : *progression free survival*

Considérations scientifiques

Pourquoi surveiller un essai clinique en cours de réalisation ?

Au cours d'un essai clinique, différents dispositifs de surveillance assurent la protection des personnes participant à l'essai et contribuent à la qualité de l'essai (Fig. 1).

Tout d'abord, l'unité de vigilance des essais examine les événements indésirables. Cette surveillance, qui relève d'une expertise de pharmacovigilance ou de matériovigilance, repose principalement sur l'analyse des données individuelles et vise à apprécier, par une méthode d'imputabilité, l'existence éventuelle d'un lien de causalité entre l'événement déclaré et l'intervention à l'étude.

Un second aspect concerne les essais de phase précoce (*e.g.*, première administration chez l'homme, escalade de dose) pour lesquels un comité d'experts analyse les données de sécurité de manière à autoriser l'inclusion d'un nouveau patient et/ou passer au palier de dose suivant. La composition de ce comité d'experts est discutée dans l'encadré 1 « Phases précoces et CSI ».

Enfin, le troisième aspect consiste à analyser des données accumulées en cours d'étude afin de s'assurer que l'incertitude quant aux avantages et inconvénients du traitement persiste, justifiant ainsi la nécessité d'acquérir de nouvelles données et donc la nécessité de poursuivre l'étude. Afin de répondre à cette question, il est nécessaire de comparer la fréquence d'événements entre les deux bras de traitement de l'étude. Mais de façon contre-productive, la connaissance et la prise en compte de ces résultats intermédiaires dans la conduite de l'étude peut, dans certaines conditions, compromettre l'intégrité des résultats. Pour cette raison, les données comparatives sont évaluées par un groupe d'experts indépendants du promoteur et des investigateurs. Ce groupe d'experts compose le Comité de surveillance indépendant (CSI) ou *Data Monitoring Committee* (DMC) en anglais. Le terme *Data Safety and Monitoring Board* (DSMB) fréquemment utilisé dans la littérature doit être abandonné dans la mesure où il met principalement l'accent sur la sécurité. Or, comme cela sera détaillé par la suite, l'analyse de la sécurité doit principalement être appréhendée comme une surveillance de la balance bénéfice-risque. Un exemple est illustré Box 1 [1].

Box 1 : l'essai APPRAISE-2 [1] évaluant l'apixaban versus placebo en addition au traitement antiagrégant plaquettaire dans les syndromes coronariens aigus a été arrêté prématurément par le CSI après l'inclusion d'environ 7 000 patients. Le motif de cet arrêt prématuré était une balance bénéfice-risque défavorable objectivée par une augmentation des saignements majeurs – HR = 2,59 [1,50 – 4,46] – non contrebalancés par une réduction des événements ischémiques – HR = 0,95 [0,80 – 1,11].

Recommandation 1

Au cours d'un essai, la protection des personnes participant à l'étude est assurée par différents dispositifs de surveillance dont il faut identifier les missions et adapter la terminologie. Il est notamment préférable d'utiliser le terme Comité de surveillance indépendant (CSI) ou le terme anglais *Data Monitoring Board* (DMC) qui – contrairement au terme *Data Safety and Monitoring Board* (DSMB) – ne font pas uniquement référence aux données de sécurité.

Quels sont les risques inhérents aux analyses en cours d'étude ?

Afin de produire des résultats robustes, un essai confirmatoire est une expérience qui s'inscrit dans un raisonnement hypothético-déductif. La démarche hypothético-déductive consiste à soumettre une hypothèse, formulée *a priori*, à une expérience *ad hoc* qui la confrontera à la réalité. Les résultats de l'expérience permettront ensuite de réfuter ou de « confirmer » l'hypothèse initiale. Un essai conçu et conduit de façon adéquate, et dont les résultats seront non biaisés et interprétables, est un essai pour lequel la confirmation ou la réfutation de l'hypothèse initiale dépend uniquement de la réalité. Pour répondre à cette prérogative, il est indispensable que l'expérimentateur, en l'occurrence l'investigateur, ne puisse influencer le résultat en modifiant la conduite de l'étude.

Lorsque la balance bénéfice-risque d'une intervention thérapeutique est évaluée en cours d'étude, c'est-à-dire avant la fin des inclusions ou avant la fin de la période de suivi planifiée, il est possible d'influencer de manière préjudiciable la poursuite de l'étude et *in fine* le résultat. En effet, la connaissance de résultats intermédiaires peut conduire à modifier le protocole – *e.g.* hiérarchie des critères de jugement d'efficacité ou critères d'inclusion. Ces modifications reposant sur des

résultats acquis en cours d'étude ont ainsi des répercussions sur l'intégrité de l'étude et conduisent à sortir l'essai d'une démarche hypothético-déductive. La connaissance des données peut aussi modifier la prise en charge future des patients de l'étude, notamment dans les études en ouvert, ou modifier la motivation des investigateurs du fait d'un résultat temporairement non conclusif ou au contraire très en faveur du traitement à l'étude.

De plus, en raison des fluctuations aléatoires d'échantillonnage, particulièrement importantes en début d'essai lorsque la quantité d'information et/ou le nombre d'événements est faible, il est tout à fait possible que les résultats puissent être ponctuellement en faveur ou en défaveur de l'efficacité du traitement évalué. Ce principe est illustré Box 2 [2]. Dans cet exemple, l'analyse finale met en évidence une absence de différence entre le bras tifacogin et le bras placebo. Pourtant, l'analyse rétrospective se limitant aux 722 premiers patients met en évidence un résultat ponctuellement en faveur d'un effet du traitement. La connaissance de ces résultats en cours d'étude offre la possibilité d'interrompre les inclusions et de conclure à l'efficacité du traitement à un moment où le résultat observé n'est pas le reflet de la réalité. D'une manière générale, le risque de conclure à tort à l'intérêt du traitement augmente avec le nombre de comparaisons ; cette augmentation du risque de conclure à tort du fait de la multiplicité des tests réalisés est communément appelée inflation du risque alpha global. Ces résultats faussement positifs sont préjudiciables pour les patients puisqu'ils peuvent conduire à l'enregistrement puis à l'utilisation de traitements sans intérêt ou de traitements qui ne sont pas supérieurs au traitement standard en cas de comparaison à ce dernier.

Box 2 : l'essai OPTIMIST [2] évaluait le tifacogin, un recombinant humain de l'inhibiteur de la voie du facteur tissulaire dans le sepsis sévère, sur la mortalité à J28.

Analyse rétrospective se limitant aux **722 premiers patients inclus** :

- 38,9 % dans le bras placebo
- 29,1 % dans le bras tifacogin ($P = 0,006$)

Analyse sur les **1754 patients inclus** :

- 33,9 % dans le bras placebo
- 34,2 % dans le bras tifacogin ($P = 0,88$)

Un arrêt après l'inclusion des 722 premiers patients aurait conclu à tort à l'intérêt du traitement.

Il faut également garder à l'esprit que cette marche aléatoire – résultats ponctuellement en faveur ou en défaveur de l'effet traitement – affecte tous les critères de jugement étudiés, y compris ceux correspondant à la sécurité du traitement. Ainsi des résultats suggérant un surrisque pourront également être observés avec un traitement n'ayant aucun défaut de sécurité, et cela d'autant plus que le nombre d'évènements indésirables analysés est important (inflation du risque alpha global de l'essai liée à la multiplicité des comparaisons).

Pour préserver l'intégrité scientifique de l'étude, il faut donc se donner les moyens de surveiller les données de sécurité et/ou d'efficacité qui s'accumulent en cours d'étude mais sans rompre la démarche hypothético-déductive et sans augmenter le risque de faux positif par inflation du risque alpha.

Quelles solutions pour préserver l'intégrité scientifique de l'étude ?

Pour permettre la surveillance de critère de sécurité et/ou d'efficacité en cours d'étude tout en préservant l'intégrité de l'étude, les données doivent donc être analysées (i) de façon confidentielle – *i.e.*, diffusées à un nombre restreint de personnes – et (ii) par des personnes indépendantes – *i.e.*, non impliquées dans le pilotage de l'essai (investigateurs, comité directeur, sponsor, promoteur, etc.). C'est pour répondre à ces prérogatives – confidentialité et indépendance – que les missions et l'organisation des CSI ont été conçues.

En effet, le CSI d'une étude doit être composé d'experts indépendants de l'étude – impliqués dans aucun autre comité (Fig. 2) – et suivre des procédures opératoires permettant de garantir l'étanchéité absolue entre le système de production et de circulation des résultats intermédiaires et les organes de pilotage de l'essai. Une telle séparation permet également au promoteur de proposer des modifications au protocole sans être suspecté de le faire du fait de la connaissance des résultats intermédiaires. Cet aspect est important car si le protocole est amendé en cours d'étude, il faut être en mesure de garantir que les modifications n'ont pas été réalisées pour s'adapter aux données à partir des tendances perçues lors une analyse intermédiaire. Les membres du CSI doivent d'ailleurs formuler au comité directeur (*Steering Committee*) et au promoteur des recommandations d'arrêt ou de poursuite de l'étude en veillant à ce qu'elles soient brèves et formulées sans divulguer la tendance prises par les résultats.

Recommandation 2

Le circuit de production et de prise en considération des résultats des analyses intermédiaires doit être distinct des organes de pilotage. Les données doivent donc être analysées par un statisticien indépendant de l'étude qui les transmet de façon confidentielle aux membres du CSI. Le CSI analyse les données et établit une recommandation quant à la poursuite, la modification ou l'arrêt de l'étude sans révéler l'orientation des résultats.

La mise en place d'un CSI est-elle obligatoire ?

La constitution d'un CSI ne doit pas être systématique et encore moins obligatoire. En réalité, plus nombreuses sont les situations où l'essai clinique ne justifie pas la constitution d'un CSI [3]. En l'absence de règle universelle codifiant la nécessité d'un CSI, il est essentiel pour le promoteur, en collaboration avec le comité directeur, de s'interroger sur la pertinence d'un CSI pour chaque essai clinique. Cette réflexion doit être engagée au moment de la planification de l'étude et les éléments à prendre en compte sont essentiellement : la population à l'étude, le type d'intervention, ainsi que le design de l'étude. Il pourra ainsi être recommandé d'ajouter une surveillance par un CSI dans les situations suivantes : maladie engageant le pronostic vital, populations vulnérables, risque d'effets indésirables important ou inconnu, règles d'arrêt reposant sur l'analyse des résultats [4]. En conformité avec l'arrêté du 2 décembre 2016, les « *motifs de constitution ou non d'un comité de surveillance indépendant* » devront être présentés au comité de protection des personnes [5].

Recommandation 3

Il n'existe pas de règle universelle pour établir la nécessité d'un CSI. Cela doit être examiné au moment de la phase de conception en fonction de la population à l'étude, du type d'intervention et du design de l'étude. L'absence de CSI doit être justifiée.

Encadré 1 : Phases précoces et CSI

La mise en place de CSI est fortement recommandée par la *Food and Drug Administration* (FDA) et l'*European Medicines Agency* (EMA) pour les essais confirmatoires (phases III). Depuis la fin des années 1990, le *National Institute of Health* (NIH) exige même un CSI pour tous les essais

confirmatoires qu'il promet. En revanche, la recommandation des agences est moins claire sur le besoin d'un CSI concernant les essais de phases précoces.

Les essais de phase I, qui ont à l'origine pour objectifs l'évaluation de la sécurité, la recherche de dose et l'évaluation des propriétés pharmacocinétiques du médicament, s'appuient sur des comités de surveillance chargés d'évaluer les événements pouvant représenter des toxicités dose-limitantes et de statuer sur l'escalade de dose. Contrairement aux CSI des essais confirmatoires, ces comités – souvent appelés *Safety Review Committee* – peuvent être composés de membres qui ne sont pas tous indépendants du promoteur et de l'investigateur, et les comités de pilotage de ces essais précoces comportent souvent des membres extérieurs [6]. La mise en place d'un CSI est également inhabituelle pour les essais de phase II, dont l'objectif est d'obtenir des données préliminaires sur l'efficacité du traitement.

Toutefois, les frontières entre les différentes phases du développement clinique d'un candidat médicament ont tendance à devenir de plus en plus confuses, les essais actuels étant souvent conçus pour répondre à plusieurs objectifs. C'est le cas des essais combinés de phase I/II, qui associent fréquemment une phase d'escalade de dose avec une ou plusieurs cohortes d'extension, randomisées ou non, ayant pour objectif de fournir des données préliminaires d'efficacité. Dans certains cas, notamment en oncologie, ces essais peuvent faire office d'études pivots [7]. Qu'il s'agisse d'essais combinés de phases I/II ou d'essais de phase II adaptatifs, des analyses intermédiaires permettent d'obtenir des résultats d'efficacité très tôt dans le plan développement. Or la variabilité de ces résultats précoces expose à un risque accru de conclure trop tôt, et à tort, au bénéfice d'un traitement. En l'absence d'une expertise externe et indépendante du comité de pilotage de l'essai et du promoteur, le risque est donc d'interrompre prématurément un essai qui se serait avéré négatif s'il avait été poursuivi. Pour les essais de phase précoce, ce risque est acceptable quand l'autorisation de mise sur le marché est conditionnée à la confirmation du bénéfice par un essai de phase III. Mais dans un contexte où des essais de phase précoce jouent le rôle d'études pivots, la présence d'un CSI se justifie pleinement [8].

Les aspects organisationnels des CSI

Une expertise nécessaire pour être membre d'un CSI

Cloisonner le circuit de production et de prise en considération des résultats est nécessaire mais reste insuffisant. Par rapport aux personnes impliquées dans le pilotage de l'essai, les membres du

CSI ont certes moins d'intérêt direct à ce que l'essai produise un résultat favorable. Pour autant, la marche aléatoire des résultats expose également les membres du CSI au risque de décisions inadaptées concernant l'efficacité ou sa sécurité, et donc, s'ils ne sont pas suffisamment formés, au même risque de remettre en cause l'intégrité scientifique de l'essai.

Pour éviter cet écueil, les décisions du CSI ne doivent pas être prises de façon naïve ou intuitive. Elles doivent suivre une méthodologie rigoureuse, basée sur des méthodes statistiques comparatives et appropriées. Par exemple, pour éviter l'inflation du risque alpha global induit par des tests répétés sur des critères d'efficacité, des méthodes d'ajustement du seuil de la signification sont impératives (*e.g.*, Lans et DeMets, Peto, etc.). Concernant la sécurité, aucune décision d'arrêt ne peut être prise sur une simple liste d'événements indésirables, ou sur un simple constat d'une différence importante même si elle est nominalement significative (valeur de $P < 0,05$) ou si elle concerne la mortalité (qui est tout aussi sujette aux fluctuations aléatoires d'échantillonnage que les autres critères). La problématique de l'inflation du risque alpha sur la multiplicité des critères de sécurité reste entière et rend complexe la tâche du CSI.

Le CSI doit donc être sensibilisé à la notion de fluctuations aléatoires d'échantillonnage, et ne pas prendre de décisions trop hâtives – qui rendraient l'essai non concluant et infructueux pour répondre à la question initiale – ou trop tardives – qui exposeraient les participants des interventions potentiellement inefficaces voire nocives [9]. Il est pour cela nécessaire de mettre en place un programme de formation spécifique. A l'image des formations validantes « bonnes pratiques cliniques [BPC] », la validation de cette formation pourra être recommandée voire rendue obligatoire pour les membres composant le CSI.

Recommandation 4

Les membres du CSI doivent être sensibilisés à l'obligation de confidentialité vis-à-vis des résultats, aux problématiques de l'analyse multiple et aux responsabilités qu'ils prennent au moment d'établir leurs recommandations. Il est pour cela nécessaire de mettre en place un programme de formation spécifique.

Composition du CSI

Si la nécessité de constituer un CSI est confirmée, le nombre est idéalement de trois ou cinq membres.

Au-delà de cinq membres, les difficultés de planification peuvent freiner le bon fonctionnement du CSI. Le calendrier des réunions tient compte du nombre de sujets à inclure, des analyses intermédiaires programmées et des échéances de l'étude. Les membres doivent s'engager pour toute la durée de l'étude. Selon les besoins de l'étude, les experts sont généralement des cliniciens accompagnés par au moins un biostatisticien ou méthodologiste ou expert en développement clinique et en analyse de données. Les membres du CSI doivent avoir une expérience dans le domaine médical étudié, y compris en recherche clinique et être indépendant de l'étude. L'indépendance de l'expert signifie qu'il n'a rien à perdre ou à gagner dans la recommandation de poursuite ou non de l'étude, y compris sur le plan personnel s'il est engagé dans un essai similaire. Idéalement, il est donc préférable d'éviter de prendre des membres engagés dans des essais concurrents.

Recommandation 5

Les membres du CSI doivent être indépendants de la conduite de l'étude, du promoteur et n'avoir aucun intérêt dans la poursuite ou l'arrêt de l'étude.

Données nécessaires pour établir les recommandations

Le CSI doit disposer, au moment de la réunion, de l'ensemble des données – idéalement monitorées – permettant de prendre une décision. Le promoteur doit tenir compte du calendrier des réunions programmées du CSI pour anticiper les actions conduisant à la production de données fiables.

Les données doivent être présentées sous forme agrégée par bras de traitement, en respectant l'anonymat des bras autant que la procédure d'investigation le permet. Si les données sont incomplètes, leur présentation doit permettre d'évaluer leur représentativité.

La fiabilité des données repose à la fois sur la qualité des données recueillies et sur l'analyse effectuée, qui doit être présentée clairement, sans ambiguïté sur l'interprétation, sous forme d'un rapport organisé selon les objectifs principaux et secondaires.

Recommandation 6

Pour pouvoir établir des recommandations fiables, le CSI doit disposer de données fiables. Le promoteur doit pour cela (i) tout mettre en œuvre pour que toutes les données nécessaires au CSI soient à jour et de qualité et (ii) respecter le calendrier de surveillance.

Les différentes missions qui peuvent être confiées au CSI

Un CSI peut être constitué pour établir des recommandations d'arrêt ou de poursuite de l'étude du fait des données de sécurité, d'efficacité et/ou de futilité, en fonction des missions qui lui sont confiées dans le protocole et dans une charte établie avant le début de l'étude.

Recommandations d'arrêt pour défaut de sécurité

L'analyse de la sécurité doit être conçue comme une surveillance de la balance bénéfice-risque. En effet, un surrisque d'effets indésirables graves peut être considéré comme acceptable si le nouveau traitement a une efficacité qui se traduit par un bénéfice clinique significatif pour les patients. C'est le cas par exemple des nouveaux anticoagulants qui apportent un bénéfice accru par rapport aux précédentes générations mais induisent également plus d'hémorragies majeures (Box 3) [10,11]. Dans des situations semblables, une surveillance de la sécurité de façon indépendante du bénéfice conduirait à un arrêt prématuré des essais pour défaut de sécurité, avec comme conséquence de priver les patients de traitements qui apportent pourtant un bénéfice clinique net supérieur aux traitements existants.

Cette évaluation nécessite de pouvoir comparer efficacité et sécurité sur des événements de même gravité clinique. Pour cela, les essais définissent un critère de jugement principal de sécurité (*safety primary endpoint*) qui est choisi pour avoir une gravité clinique identique aux critères de jugement principal d'efficacité. Dans d'autres cas, un critère de bénéfice clinique net est utilisé comme critère de jugement principal. La mortalité toute cause permet aussi de juger du bénéfice clinique net quand les effets indésirables sont potentiellement mortels.

Box 3 : L'essai TRITON-TIMI 38 [10] a évalué le prasugrel, un nouvel antiagrégant plaquettaire, versus le clopidogrel dans le syndrome coronarien aigu. Une augmentation relative de 32 % des saignements majeurs a été observée – HR = 1,32 [1,03 – 1,68] –, avec cependant une réduction relative des événements ischémiques de 19 % – HR = 0,81 [0,73 – 0,90].

L'essai ISIS-2 [11] a évalué à l'aide d'un plan factoriel la streptokinase et/ou l'aspirine à la phase aiguë du syndrome coronarien. Sept hémorragies cérébrales ont été observées dans le groupe streptokinase versus aucune dans le groupe placebo mais cette augmentation d'hémorragies cérébrales s'accompagne d'une réduction importante de la mortalité d'origine vasculaire à 5 semaines – OR = 0,75 [0,68 – 0,83].

Une difficulté à prendre en compte dans l'évaluation de la balance bénéfice-risque est la fréquente dissociation temporelle entre (i) l'apparition des événements indésirables qui peuvent survenir rapidement après l'initiation du traitement – que ce soit une chirurgie, un dispositif implantable ou un médicament – et (ii) le temps nécessaire pour percevoir le bénéfice clinique. Par exemple, pour un traitement de première ligne en oncologie, les effets indésirables peuvent survenir pendant l'administration du traitement, soit en début d'essai. Mais la démonstration du bénéfice en termes de survie nécessitera un suivi de plusieurs mois voire plusieurs années, puisque les décès surviennent généralement après plusieurs progressions et plusieurs lignes de traitements.

En plus de ces difficultés temporelles, s'ajoute le principe de précaution qui consiste à considérer un défaut de sécurité sur la base d'une simple suspicion, sans besoin d'en avoir la preuve formelle. Il est pourtant indispensable de garder à l'esprit que la surveillance de la sécurité est rendue délicate par le risque de « faux positifs » dû (i) aux fluctuations aléatoires d'échantillonnage et (ii) à la multiplicité des comparaisons. Concernant les fluctuations aléatoires d'échantillonnage, elles sont d'autant plus importantes que le nombre de patients est faible – début de l'essai – ou que la fréquence de l'événement observé est faible (Box 4) [12,13]. Les « fausses différences » qui découlent de ces fluctuations peuvent amener le CSI à conclure à tort à un surrisque d'événements. Concernant la multiplicité des comparaisons, elle est due au fait que de nombreux types d'événements indésirables voire différents regroupements d'événements sont comparés, ce qui va également contribuer à augmenter le risque de « faux positif ».

Box 4 : l'essai CANVAS a évalué la canagliflozine versus placebo dans le diabète de type 2.

L'objectif principal était la non-infériorité sur les événements cardiovasculaires. Au cours d'une analyse intermédiaire, le CSI a identifié un nombre plus important d'amputations dans les groupes

canagliflozine : 7 patients sur 1 000 pour les patients du groupe canagliflozine 100 mg par jour ; 5 patients sur 1 000 pour les patients du groupe canagliflozine 300 mg par jour ; 3 patients sur 1 000 pour les patients du groupe placebo [12]. Ce risque n'était pas retrouvé dans un autre essai similaire – l'essai CANVAS-R – en cours également. La décision du CSI a été de continuer l'essai, mais les amputations avaient été ajoutées par les agences à la liste des effets indésirables de la canagliflozine. Après cette analyse intermédiaire, les données de sécurité des essais CANVAS et CREDENCE – réalisés dans la néphropathie diabétique –, ont été regroupées et après analyse approfondie des données, il a été conclu que le surrisque d'amputation initialement observé dans l'essai CANVAS était certainement dû au hasard [13].

Recommandation 7

En cas de recommandation d'arrêt pour sécurité, le CSI doit se s'assurer que les données de l'essai sont suffisamment convaincantes pour conclure à une balance bénéfice-risque non favorable.

Recommandations d'arrêt pour efficacité

De prime abord, la surveillance de l'efficacité peut sembler moins délicate et moins subjective que la surveillance de la sécurité car elle repose sur des règles statistiques prévues au protocole (Lan et DeMets, Peto-Haybitte, etc.) ayant pour objectif de maîtriser l'inflation du risque alpha. Pour rappel, ces règles d'arrêt doivent impérativement être fixées a priori (Box 5) [14].

Box 5 : l'essai FAME 2 [14] évaluait la mesure de la réserve coronarienne pour guider la revascularisation coronarienne chez des patients ayant un angor stable. Le résumé relate que l'essai a été interrompu prématurément : « *Recruitment was halted prematurely after enrollment of 1220 patients (888 who underwent randomization and 332 enrolled in the registry) because of a significant between-group difference in the percentage of patients who had a primary end-point event* ». Pourtant le protocole précisait qu'aucune analyse intermédiaire n'était planifiée « *Interim analysis : no interim analysis is planned* » (page 26 section 13.3.2).

La justification de l'arrêt prématuré est l'existence d'une différence significative. Mais cette différence n'aurait jamais dû être connue de l'investigateur principal étant donné l'absence de méthode de prise en charge de la multiplicité engendrée par d'éventuelles analyses intermédiaires. Ainsi cette décision d'arrêt repose sur la connaissance des résultats et rien ne permet d'écarter qu'il s'agisse d'une simple fluctuation aléatoire particulièrement favorable, mais ne reflétant pas la réalité de l'efficacité du traitement.

Initialement, les analyses intermédiaires d'efficacité étaient utilisées de manière à arrêter précocement un essai – *i.e.*, avant l'inclusion de la totalité des patients – dès lors que la démonstration d'efficacité était obtenue sur le critère de jugement principal. Mais, il est utile de rappeler qu'un arrêt trop prématuré est susceptible de compromettre l'interprétation de la balance bénéfice-risque. Ainsi, malgré la démonstration d'efficacité sur le critère de jugement principal, il peut être nécessaire de poursuivre les inclusions si un doute persiste quant à la balance bénéfice-risque (Box 6, Fig. 3.) [15].

Box 6 : l'essai COMPASS [15] – dont le design est présenté Fig. 3 – a comparé l'intérêt du rivaroxaban ou de son association à dose réduite à l'aspirine par rapport à l'aspirine seule en prévention cardiovasculaire secondaire. Le plan de contrôle du risque alpha global aménageait la possibilité d'apporter des démonstrations sur plusieurs critères de jugement dont la mortalité toute cause.

Lors de la première analyse intermédiaire, le CSI a recommandé l'arrêt de l'étude pour efficacité devant une différence sur le premier critère de jugement de la hiérarchie. Cette différence sur les événements cardiovasculaires était observée uniquement dans le bras rivaroxaban en association à l'aspirine versus aspirine seule – HR = 0,76 [0,66 – 0,86] ; $P < 0,001$. Malgré une absence de différence entre le bras du rivaroxaban seul et le bras aspirine seule – HR = 0,90 [0,78 – 1,03] ; $P = 0,12$ –, les inclusions dans le bras rivaroxaban ont également été arrêtées.

Tel que planifié, le fait que la différence entre le rivaroxaban et l'aspirine soit non significative pour le premier critère de jugement interrompait l'analyse et par conséquent, les critères de jugement suivants ne pouvaient être testés. « *The study protocol and Statistical Analysis Plan do not specify modifications to the Hochberg gatekeeping procedure for the testing of secondary efficacy outcomes in the case of a premature termination for efficacy when both comparisons are stopped but only one of the comparisons met the modified Haybittle-Peto boundary* ».

Au total, le seul résultat démontré – correspondant à la réduction des événements cardiovasculaires mortels et non mortels – ne permet pas de conclure à une balance bénéfice-risque favorable. En effet, cet effet sur les événements cardiovasculaires est complètement contrebalancé par une augmentation des saignements majeurs qui était le critère de jugement principal de sécurité (12 saignements majeurs de plus pour 13 événements ischémiques de moins pour 1000 patients traités 23 mois).

Ce surcroît de saignements majeurs aurait pu éventuellement être acceptable si une réduction de mortalité toute cause (quatrième critère de la hiérarchie) avait été démontrée. Mais la décision d'arrêter l'essai COMPASS précocement sur recommandations du CSI rend impossible l'analyse de ce critère de jugement. Et si une tendance semble se dégager – HR = 0,82 [0,71 – 0,96], elle doit être considérée comme non statistiquement significative et ne peut être prise en considération pour la décision. D'ailleurs, les p-values des critères de jugement secondaires ne devraient pas figurer dans le tableau des résultats puisqu'en conformité avec le plan d'analyse statistique, ces analyses n'auraient pas dû être réalisées.

Cet exemple illustre parfaitement le fait qu'une étude doit être arrêtée prématurément uniquement lorsque les résultats produits permettent de conclure de manière solide et irréfutable à une balance bénéfice-risque favorable (et non pas seulement à une certaine efficacité).

À l'aide d'un plan de contrôle du risque alpha global adapté, la démonstration d'efficacité peut désormais être apportée non pas sur un seul et unique critère de jugement mais sur plusieurs critères de jugement. Les « règles d'arrêt » peuvent alors être utilisées de manière différente : elles permettent de faire des démonstrations anticipées sur un ou plusieurs critères de jugement sans pour autant conduire à l'arrêt définitif de l'essai. L'essai est en effet poursuivi de manière à se donner les moyens de conclure sur l'ensemble des critères de jugement. La démonstration d'efficacité sur le ou les premiers critères de jugement peut éventuellement servir à déposer une demande d'autorisation de mise sur le marché auprès des autorités compétentes. Par exemple, les critères de jugement fréquemment utilisés en oncologie sont : (i) la survie sans progression (PFS [*progression free survival*]) et (ii) la survie globale (OS [*overall survival*]). Dans les études pivots, ces deux critères sont souvent hiérarchisés, avec la PFS en premier puis l'OS. L'enjeu est bien évidemment de montrer un bénéfice sur l'OS mais la maturité des données est plus longue que pour la PFS (Box 7) [16].

Box 7. L'essai RUBY [16] évalue le dostarlimab en combinaison à une chimiothérapie à base de sels de platine pour le traitement des patientes adultes atteintes de cancer de l'endomètre avancé ou récidivant. Après un suivi médian de 25 mois, les résultats de l'analyse intermédiaire sur l'ensemble des patients inclus montrent un effet sur la PFS – HR = 0,64 [0,51 – 0,80]. Les résultats sur l'OS sont par contre non significatifs et le suivi est poursuivi pour confirmer ou infirmer un effet favorable sur la mortalité. L'essai RUBY est donc toujours en cours mais compte tenu de la démonstration d'un bénéfice sur la PFS, le dostarlimab a déjà reçu une autorisation d'accès précoce dans l'indication concernée par cet essai.

Lorsque des analyses intermédiaires sont prévues, il est également important de ne pas confondre l'efficacité du nouveau traitement avec un problème de sécurité du traitement contrôle. Pour illustrer ce principe, prenons l'exemple d'un nouveau traitement réduisant la mortalité par rapport au traitement de référence. En début d'essai, la quantité d'information est insuffisante pour conclure à un bénéfice sur la mortalité : la règle d'arrêt n'est pas atteinte. Mais à ce stade, une situation où le nombre de décès serait plus important dans le groupe contrôle pourrait laisser penser que le traitement de référence entraîne un surcroît de mortalité, ce qui pourrait conduire à arrêter l'étude par précaution du fait d'un problème de sécurité dans le groupe contrôle. Pourtant, dans la mesure où il s'agit du traitement de référence, les patients randomisés dans le groupe contrôle auraient reçu le même traitement s'ils n'avaient pas participé à l'étude. La poursuite de l'étude ne constitue donc pas une perte de chance pour les participants qui seront randomisés dans le groupe contrôle. Cependant, arrêter l'essai par précaution et sur la base d'une mauvaise interprétation de la situation (i) prive de la possibilité de démontrer le bénéfice du nouveau traitement sur la mortalité et (ii) pose un problème éthique, car l'essai expose les participants aux contraintes d'une recherche qui fournira des résultats ne permettant pas d'apporter une connaissance scientifique et médicale satisfaisante. La mise en place d'autres études sera nécessaire pour apporter la réponse à la question posée et retarde d'autant l'avancée des connaissances.

En conclusion, même si cela peut paraître séduisant, il n'y a pas d'urgence à arrêter un essai pour efficacité. Si l'essai a débuté, c'est que la totalité de ce qui est prévu (en nombre de sujets à inclure et en durée de suivi) a été jugée acceptable au niveau éthique et nécessaire étant donné que le nouveau traitement représente une piste de progrès thérapeutique suffisamment prometteuse. Il serait donc dommageable de l'interrompre de façon précipitée avant qu'il fournisse les preuves robustes nécessaires pour changer les pratiques et utiliser le nouveau traitement. Un tel arrêt pose d'ailleurs un double problème éthique : celui d'avoir exposé des patients à d'éventuels risques avec

le nouveau traitement pour rien (Box 8) [17] et celui de priver les patients d'un progrès thérapeutique par insuffisance de preuve en raison de la compromission de l'intégrité scientifique de l'étude.

Box 8. L'essai LIGHT [17] a évalué la survenue d'évènements cardiovasculaires chez des patients en surpoids ou obèses traités par l'association naltrexone-bupropion versus placebo. En effet, afin d'exclure un risque accru d'évènements cardiovasculaires, la FDA avait exigé un essai de non-infériorité dont la marge de non-infériorité – correspondant à la borne supérieure de l'intervalle de confiance du HR – était fixée à 1,4. Une analyse intermédiaire était également planifiée après que 25 % des évènements attendus étaient survenus et l'étude était poursuivie uniquement si la borne supérieure de l'intervalle de confiance du HR était inférieure à 2.

Mais en raison de la publication de cette analyse intermédiaire par le sponsor, contrevenant au principe de confidentialité, il a été décidé de clore l'étude sans pouvoir évaluer la non-infériorité. Par conséquent, malgré l'inclusion de 8910 patients, l'essai LIGHT ne permet pas de répondre à la question posée par l'essai et il n'est pas possible d'exclure que l'association naltrexone-bupropion entraîne un surrisque d'évènements cardio-vasculaires.

Arrêter de façon anticipée un essai pour efficacité dans le but que les patients aient plus rapidement accès au nouveau produit est très discutable. En effet, cet arrêt expose les patients à une interruption thérapeutique, la poursuite du traitement devant se faire par un accès dérogatoire parfois complexe à mettre en place. De plus, l'accès au nouveau produit ne sera pas basé sur une seule et unique étude mais sur l'ensemble des données d'efficacité et de sécurité issues des études réalisées dans la cadre du plan de développement du nouveau produit. Cette évaluation de la balance bénéfice-risque du nouveau traitement au regard de toutes les données disponibles prend du temps, auquel s'ajoute une évaluation sur l'intérêt d'une prise en charge par la solidarité nationale puis une phase de fixation du remboursement et de fixation du prix

Un arrêt trop précoce peut au contraire se révéler contre-productif pour les patients en fragilisant la preuve de l'efficacité et de l'innocuité et donc en mettant en péril l'accès au nouveau produit en raison d'un niveau de preuve insuffisant. Certains arrêts anticipés pour efficacité sont en réalité sous tendus par des intérêts financiers et compétitifs éloignés de l'intérêt des patients [18].

En revanche, un arrêt trop précoce pour efficacité pose les difficultés suivantes qui peuvent rendre caduque le résultat, car trop fragile et peu convainquant :

- Faible durée d'exposition limitant l'appréciation de la sécurité ;
- Faible précision des estimations des effets traitement (intervalle de confiance ajusté large touchant l'absence d'effet) ;
- Surestimation de l'effet traitement [19] ;
- Taille de l'effet irréaliste ou analyse faite à une fraction d'information non réaliste ;
- Difficulté pour apprécier la cohérence des résultats entre les critères de jugement et les sous-groupes ;
- Faible précision des courbes de Kaplan-Meier avec des censures liées à la date de point très précoces ;
- Etc.

Recommandation 8

En cas de recommandation d'arrêt pour efficacité, le CSI doit se s'assurer que les données de l'essai sont suffisamment convaincantes pour en faire un standard of care.

Recommandations d'arrêt pour futilité

Quand l'essai évalue une pratique médicale devenue courante sans preuve de son bénéfice, l'arrêt pour futilité est à proscrire (le protocole ne devrait même pas le prévoir). En effet, si ce traitement n'apporte pas de bénéfice, il faut être en mesure de produire un niveau de preuve suffisamment probant pour convaincre la communauté d'arrêter une pratique médicale déjà établie. Dire que l'essai a échoué à montrer le bénéfice de l'intervention aura une portée insuffisante contrairement à (i) un résultat montrant l'absence d'effet avec un intervalle de confiance étroit ou (ii) un effet délétère. Dans cette situation l'essai doit aller à son terme, y compris si des considérations de sécurité apparaissent. En effet, dans un contexte de pratique médicale établie, il est indispensable que l'essai apporte une preuve forte du défaut de sécurité pour que la communauté renonce à cette pratique (Box 9) [20,21]. Un surrisque peu établi peut s'avérer insuffisant pour emporter pas la conviction. Ainsi, même si l'une est moins efficace ou moins sûre que l'autre, la participation à l'étude n'entraîne pas une perte de chance par rapport à la prise en charge habituelle de ces patients.

Box 9. L'essai par Combes et al., [20] comparait la mortalité à 60 jours de patients ayant un SRAS traité par ECMO versus une ventilation standard. L'essai a été arrêté par son CSI pour futilité à une fraction d'information de 73% car la limite de futilité prévue au protocole avait été franchie. Le résultat sur la mortalité était un risque relatif de 0,76 [0,55 – 1,04] ; $P = 0,09$. Ce résultat pourtant à la limite de la mise en évidence d'un bénéfice de mortalité est difficile à interpréter et fait regretter, même si les règles ont été appliquées, l'arrêt de l'étude pour futilité. Clairement il ne permet pas de conclure à une absence convaincante de bénéfice de l'ECMO mais il ne permet pas non plus de conclure à son bénéfice. Cet exemple rappelle que les règles statistiques doivent être qu'un guide pour le CSI qui doit chercher impérativement à « voir par le dessus » la totalité de la situation [21].

Recommandation 9

En cas de recommandation d'arrêt pour futilité, et dans un contexte de pratique médicale déjà établie, le CSI doit se s'assurer que les données de l'essai sont suffisamment convaincantes pour statuer sur l'absence d'intérêt de l'intervention.

Les recommandations du CSI doivent également tenir compte du facteur temps. La recommandation peut revêtir un caractère d'urgence lorsqu'un arrêt d'étude doit être recommandé sur la base de données de sécurité, pour le bien des patients inclus, alors que la recommandation pour efficacité/futilité peut être un peu temporisée.

Publication des résultats des analyses intermédiaires

D'un point de vue méthodologique, il est universellement reconnu que les résultats des analyses intermédiaires non concluantes ne doivent pas être communiqués en raison du risque de compromettre l'intégrité scientifique de l'étude. Pourtant, il n'est pas exceptionnel, notamment en oncologie, que les résultats de ces analyses intermédiaires soient publiés. Tel que décrit précédemment, les deux principaux critères de jugement utilisés en oncologie sont la survie sans progression et la survie globale. La démonstration d'un bénéfice sur ces deux critères s'obtient avec un décalage dans le temps : en effet, les événements de survie sans progression intègrent les progressions radiologiques qui surviennent plus rapidement que les décès. Les plans d'analyses

intermédiaires tiennent compte de ce décalage dans le temps et la démonstration d'un bénéfice sur la survie sans progression est obtenue avant celle d'un bénéfice sur la survie globale. Dans la mesure où la démonstration d'un bénéfice sur la survie sans progression offre la possibilité d'aller chercher un enregistrement de la nouvelle molécule auprès des autorités, ces résultats sont régulièrement communiqués en congrès et font l'objet d'une publication.

En eux-mêmes, la publication des résultats concernant la survie sans progression ne pose pas de problèmes méthodologiques car ils sont définitivement acquis. Mais généralement, ces publications rapportent également les résultats non concluants de l'analyse intermédiaire sur la survie globale. En effet, l'analyse de la survie globale est généralement faite de manière synchrone à l'analyse de la survie sans progression. Les papiers précisent alors que l'on ne peut pas tenir compte de la négativité de cette analyse d'OS, car les données ne sont pas matures. Dans d'autres cas, ces résultats peuvent conduire à un spin de conclusion si l'effet sur la survie globale est nominalelement significatif (Box 10) [22]. Bien que non concluant, le spin suggère que la démonstration sur la survie globale est d'ores et déjà obtenue.

Box 10. L'essai FIRST [22] illustre le principe du spin de conclusion : *“Although the difference in overall survival did not cross the prespecified superiority boundary ($P < 0.0096$), continuous lenalidomide–dexamethasone reduced the risk of death, as compared with MPT (hazard ratio, 0.78; 95% CI, 0.64 to 0.96; $P = 0.02$)”.*

Cette façon de communiquer des résultats non concluants, tout en suggérant une possible tendance favorable peut perturber la conduite de l'essai et remettre en cause l'intégrité de l'essai vis-à-vis de son enjeu primordial, à savoir démontrer un bénéfice en termes de survie globale. Les investigateurs – ou les patients eux-mêmes – pourront ne pas percevoir l'incertitude autour de cette tendance et vouloir sortir de l'étude les patients afin de prescrire le traitement à l'étude si celui-ci est déjà disponible. Ces sorties prématurées exposent alors au risque de ne pas pouvoir conclure sur ce critère d'intérêt et l'incertitude quant au bénéfice du traitement sur la mortalité globale persistera.

Recommandation 10

Lorsque le plan d'analyse statistiques prévoit une analyse hiérarchisée des critères de jugement, seuls les résultats des analyses concluantes doivent être communiqués. En effet, les résultats des critères non concluants et pour lesquels l'étude est poursuivie ne doivent pas être communiqués.

Attractivité de la fonction de membre d'un CSI

Les membres de CSI contribuent à la sécurité des participants des essais cliniques et sont les garants de l'intégrité scientifique de ces recherches. Compte-tenu de la responsabilité scientifique et médicale, de l'expertise et du temps nécessaires, la fonction de membre d'un CSI est une mission d'intérêt général essentielle qui devrait être reconnue et valorisée.

Les membres de CSI sont souvent des professionnels de santé hospitaliers, hospitalo-universitaires ou universitaires dont la charge de travail est importante. La participation à des CSI pourrait être reconnue en valorisant l'évolution de carrière (échelons, primes).

- Reconnaissance professionnelle : hospitalière, académique par l'attribution d'un temps dédié à la recherche clinique ;
- Indemnisation financière : possible ;
- Publication : les membres du CSI doivent être informés qu'ils ne feront pas partie des auteurs de la publication.

Recommandation 11

Le temps et l'expertise liés à la participation à des CSI (des essais à promotion institutionnelle non rémunérés) devraient être valorisés au titre des activités collectives d'intérêt général (comme c'est le cas pour la participation comme membre ou président de CPP, ou comme expert évaluateur des produits de santé (Agence nationale de sécurité du médicament et des produits de santé [ANSM], EMA, Haute autorité de santé [HAS], etc.) et pris en compte dans les de promotion hospitalière ou hospitalo-universitaire.

En conclusion, les CSI contribuent à la protection des personnes participant à l'essai ; mais pour autant, tous les essais cliniques ne nécessitent pas la mise en place d'un CSI. Le recours ou non à un CIS doit donc être anticipé et justifié dès la phase de conception de l'étude. Dès lors que l'essai

nécessite un CSI, la surveillance des données doit être réalisée avec la plus grande rigueur méthodologique pour ne pas risquer de remettre en question l'intégrité scientifique de l'étude et la portée des résultats en compromettant leur degré de crédibilité. En amont des réunions, le promoteur devra tout mettre en œuvre pour que les données nécessaires à l'élaboration de recommandations soient à jour et de qualité. De leur côté, les membres des CSI doivent être conscients de leur responsabilité et de la complexité de la mission qui leur ait confié. En effet, leurs décisions doivent intégrer les problématiques méthodologiques liées au respect de la démarche hypothético-déductive mais aussi celle liées aux aspects statistiques des analyses répétées. Ces prises de décision demandent une forte expertise des membres du CSI et le respect de toutes les règles d'organisation logistique des analyses intermédiaires. Pour cette raison, il est nécessaire que les membres des CIS soient parfaitement formés aux enjeux scientifiques qui doivent guider leurs recommandations.

Déclaration de liens d'intérêts

Les auteurs n'ont pas de conflits d'intérêts à déclarer

Références

- [1] Alexander JH, Lopes RD, James S, Kilaru R, He Y, Mohan P, et al. Apixaban with antiplatelet therapy after acute coronary syndrome. *N Engl J Med* 2011;365:699–708.
- [2] Abraham E, Reinhart K, Opal S, Demeyer I, Doig C, Rodriguez AL, et al. Efficacy and safety of tifacogin (recombinant tissue factor pathway inhibitor) in severe sepsis: a randomized controlled trial. *JAMA* 2003;290:238–47.
- [3] EMA. Data monitoring committees - Scientific guideline. European Medicines Agency 2018. <https://www.ema.europa.eu/en/data-monitoring-committees-scientific-guideline>. [Consulté le 27 octobre 2023].
- [4] Evans SR. Independent oversight of clinical trials through data and safety monitoring boards. *NEJM Evidence* 2022;1:EVIDctw2100005. <https://evidence.nejm.org/doi/full/10.1056/EVIDctw2100005>. [Consulté le 27 octobre 2023].
- [5] Legifrance. Arrêté du 2 décembre 2016 fixant le contenu, le format et les modalités de présentation du dossier de demande d'avis au comité de protection des personnes sur un projet de recherche mentionnée au 1° de l'article L. 1121-1 du code de la santé publique portant sur un médicament à usage humain. Décembre 2016. <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000033545400>. [Consulté le 27 octobre 2023].
- [6] EMA. Data Monitoring Committees issues - Scientific guideline. European Medicines Agency 2018. <https://www.ema.europa.eu/en/data-monitoring-committees-issues-scientific-guideline>. [Consulté le 27 octobre 2023].
- [7] Adashek JJ, LoRusso PM, Hong DS, Kurzrock R. Phase I trials as valid therapeutic options for patients with cancer. *Nat Rev Clin Oncol* 2019;16:773–8.
- [8] Van Norman GA. Data safety and monitoring boards should be required for both early- and late-phase clinical trials. *JACC Basic Transl Sci* 2021;6:887–96.
- [9] Schulman S, Carlsson A, Laporte S. When is it safe to stop for safety? *J Thromb Haemost.* 2012;10:1188-90.
- [10] Wiviott SD, Braunwald E, McCabe CH, Montalescot G, Ruzyllo W, Gottlieb S, et al. Prasugrel versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* 2007;357:2001–15.
- [11] ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;332:349–60.
- [12] Food and Drug Administration. Research C for DE and. FDA Drug Safety Podcast: Interim clinical trial results find increased risk of leg and foot amputations, mostly affecting the toes, with

the diabetes medicine canagliflozin (Invokana, Invokamet); FDA to investigate. 2022.

<https://www.fda.gov/drugs/fda-drug-safety-podcasts/fda-drug-safety-podcast-interim-clinical-trial-results-find-increased-risk-leg-and-foot-amputations>. [Consulté le 27 octobre 2023].

[13] Arnott C, Huang Y, Neuen BL, Di Tanna GL, Cannon CP, Oh R, et al. The effect of canagliflozin on amputation risk in the CANVAS program and the CREDENCE trial. *Diabetes Obes Metab* 2020;22:1753–66.

[14] De Bruyne B, Pijls NHJ, Kalesan B, Barbato E, Tonino PAL, Piroth Z, et al. Fractional flow reserve–guided PCI versus medical therapy in stable coronary disease. *N Engl J Med* 2012;367:991–1001.

[15] Eikelboom JW, Connolly SJ, Bosch J, Dagenais GR, Hart RG, Shestakovska O, et al. Rivaroxaban with or without aspirin in stable cardiovascular disease. *N Engl J Med* 2017;377:1319–30.

[16] Mirza MR, Chase DM, Slomovitz BM, dePont Christensen R, Novák Z, Black D, et al. Dostarlimab for primary advanced or recurrent endometrial cancer. *N Engl J Med* 2023;388:2145–58.

[17] Nissen SE, Wolski KE, Prcela L, Wadden T, Buse JB, Bakris G, et al. Effect of naltrexone-bupropion on major adverse cardiovascular events in overweight and obese patients with cardiovascular risk factors: a randomized clinical trial. *JAMA* 2016;315:990–1004.

[18] Iltis AS. Stopping trials early for commercial reasons: the risk–benefit relationship as a moral compass. *J Med Ethics* 2005;31:410–4.

[19] Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA* 2010;303:1180–7.

[20] Combes A, Hajage D, Capellier G, Demoule A, Lavoué S, Guervilly C, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. *N Engl J Med* 2018;378:1965–75.

[21] Harrington D, Drazen JM. Learning from a trial stopped by a data and safety monitoring board. *N Engl J Med* 2018;378:2031–2.

[22] Benboubker L, Dimopoulos MA, Dispenzieri A, Catalano J, Belch AR, Cavo M, et al. Lenalidomide and dexamethasone in transplant-ineligible patients with myeloma. *N Engl J Med* 2014;371:906–17.

Légendes des figures



Figure 1. Dispositifs de surveillance assurant la protection des personnes en cours d'étude.

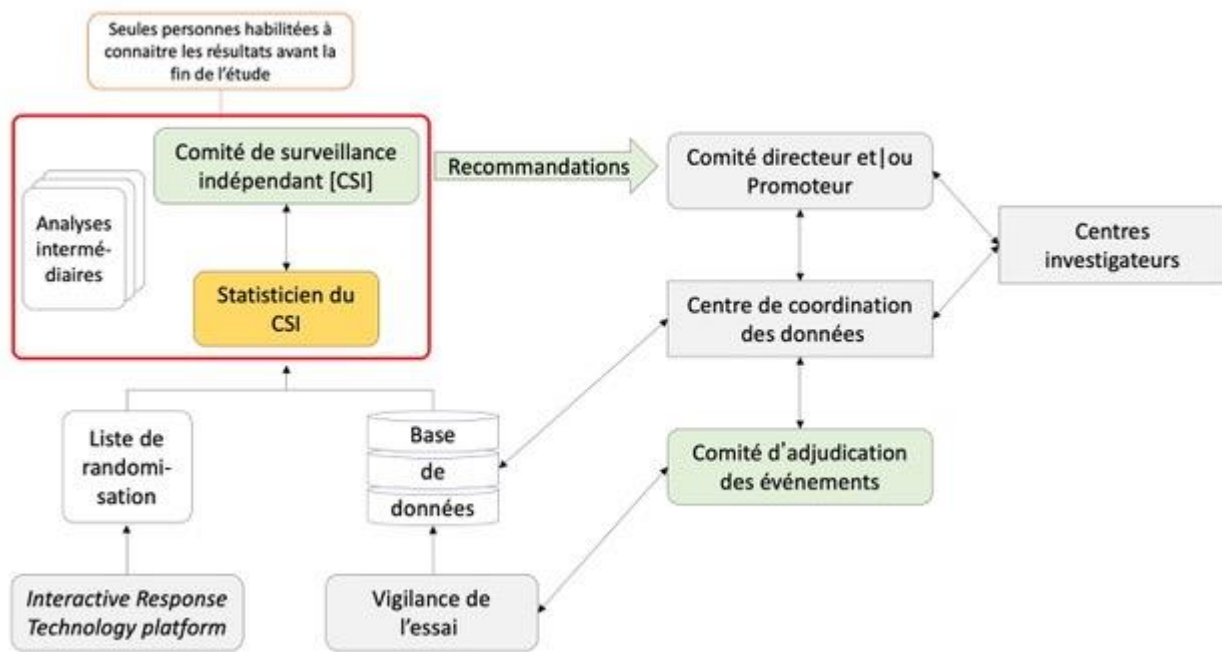


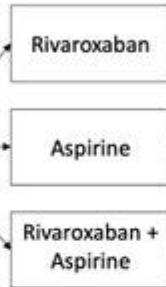
Figure 2. Comités et acteurs dans un essai. Les membres du Comité de surveillance indépendant (CSI) doivent être indépendants et à ce titre, ils ne doivent être impliqués dans aucun autre comité que le CSI.

Essai COMPASS

Patients éligibles

- Maladie coronarienne
- Maladie artérielle périphérique
- ou le deux

Randomisés
1:1:1
N = 27 395 patients



4 critères de jugement hiérarchisés	CJ 1	CJ 2	CJ 3	CJ 4
Accident vasculaire cérébral	✓			
Accident vasculaire cérébral ischémique		✓	✓	
Infarctus du myocarde	✓	✓	✓	
Ischémie aiguë des membres		✓	✓	
Décès d'origine cardiovasculaire	✓			
Décès dû à une maladie coronarienne		✓	✓	
Décès toutes causes				✓

Multiplicité des tests liée

- aux deux bras de traitements (rivaroxaban et rivaroxaban + aspirine) à comparer à l'aspirine seule
 - aux quatre critères de jugement hiérarchisés
 - aux analyses intermédiaires
- prise en compte par un plan de contrôle du risque alpha global

Figure 3. Design de l'essai COMPASS. La stratégie d'analyse prévoyait de tester quatre critères de jugement hiérarchisés en comparant les deux modalités d'utilisation du rivaroxaban (seul ou en association à l'aspirine) à l'aspirine en monothérapie.