

Le principe de l'intention-to-treat

Peter Kleist

GlaxoSmithKline AG, Münchenbuchsee

Quintessence

- Les essais cliniques doivent a priori être interprétés selon le principe de l'intention-to-treat (ITT). L'évaluation des données doit porter sur tous les patients randomisés en conservant leur attribution aux différents groupes et indépendamment du déroulement de l'étude.
- L'ITT ne minimise pas le biais résultant de l'absence de certaines données relatives aux end-points de l'étude. Il est essentiel de continuer à suivre tous les patients et à limiter au maximum les données manquantes si on veut que la qualité des interprétations ITT soit garantie. Les méthodes statistiques pour traiter les données manquantes ne sont pas exemptes de problèmes.
- Malgré ses limites, l'ITT reste le meilleur moyen dont nous disposons pour éviter les biais d'interprétation et pour évaluer les effets thérapeutiques dans le cadre des essais cliniques.

Le problème

Pas plus que la vie ne suit fidèlement les manuels, les essais cliniques ne se déroulent jamais de façon idéale. On peut ainsi être confronté dès le départ au non-respect de certains critères de sélection fixés par le protocole, conduisant à l'inclusion erronée de certains patients. Et comme les patients suivent dans le fond leurs propres règles, ils prennent souvent les médicaments de l'essai de manière irrégulière, parfois même pas du tout. Il peut aussi arriver qu'ils interrompent le traitement prématurément, déménagent dans une autre ville ou décèdent avant que l'étude ne soit terminée.

Alors comment tenir compte de cette réalité? Faut-il exclure ces cas de l'interprétation finale et simplement ignorer ces données? Quelle est la méthode d'analyse appropriée?

Dans le domaine scientifique et chez les autorités de contrôle des médicaments, le principe de l'intention-to-treat (ITT) s'est entre-temps établi comme règle d'or de l'analyse primaire, autrement dit comme l'analyse statistique validée pour les essais cliniques [1, 2]. Nous allons expliquer dans cet article ce qu'est le principe de l'ITT et le rôle qu'il joue dans les essais contrôlés et randomisés.



Peter Kleist

L'auteur certifie qu'aucun conflit d'intérêt n'est lié à cet article.

Définition et délimitation

Le principe de l'intention-to-treat dit que *tous* les patients randomisés doivent être pris en compte par

l'analyse – et ceci en maintenant leur répartition dans les différents groupes de l'essai par la randomisation, indépendamment de leur devenir au cours de l'étude. La notion «intention-to-treat» décrit le groupe complet des patients que l'on avait l'intention de traiter au moment de la randomisation. Des patients inclus par erreur, ainsi que les patients ayant interrompu le traitement ou l'étude en cours de route entrent par conséquent aussi dans l'analyse. Nos collègues anglais qualifient cette approche de façon très explicite: «Once randomised, always analysed».

La valeur d'une évaluation ITT est néanmoins dépendante de la dimension de l'exhaustivité des informations concernant les patients. L'ITT ne commence par conséquent pas seulement lors de l'analyse des données, mais elle englobe bien la totalité de l'étude. Les protocoles doivent être conçus d'emblée de telle façon que les données manquantes soient l'exception et non la règle. Tous les patients doivent être observés dans toute la mesure du possible jusqu'au terme prévu de l'essai, même lorsque le traitement a été interrompu prématurément, par exemple à la suite d'effets indésirables. La minimisation du taux de pertes de patients (loss-to-follow-up) est absolument essentielle [3]. Ce n'est que la saisie complète des données qui permettra finalement une évaluation sans distorsion de l'efficacité et de la sécurité d'une mesure thérapeutique donnée. Comme l'ITT constitue un concept global, qui caractérise un essai dans son entier, et pas simplement

Le principe de l'intention-to-treat dit que tous les patients randomisés doivent être pris en compte par l'analyse – «Once randomised, always analysed»

une méthode d'analyse appliquée à la fin du protocole, nous considérons l'ITT comme un principe. Contrairement à l'analyse ITT, l'analyse «per-protocole» ne porte que sur les patients qui ne se sont pas écartés sensiblement des prescriptions du protocole. Les violations du protocole, comme décrites précédemment, conduiraient ici à l'exclusion des patients concernés de l'analyse statistique. Il devient évident que la non-prise en compte des patients et de leurs données lors d'une évaluation ITT planifiée revient à une analyse «pseudo-ITT», en fait proche d'une évaluation per-protocole [4].

Une troisième méthode d'évaluation est celle de l'analyse «as-treated» (telle que traité dans la réalité); elle joue un rôle lorsqu'une proportion significative des patients ne reçoit pas le traitement attribué par la randomisation,

mais un autre traitement. Les analyses as-treated s'utilisent surtout dans les essais comparatifs avec plusieurs médicaments ou interventions chirurgicales. Elles diffèrent des analyses ITT et per-protocole par le fait qu'on fait abstraction de la randomisation initiale.

Pourquoi le principe de l'ITT?

Les arguments suivants parlent en faveur du principe de l'ITT comme méthode privilégiée d'évaluation des résultats d'essais cliniques:

- La comparabilité entre les différents groupes de patients est maintenue, car tous les patients randomisés sont inclus.
- La puissance du test de l'essai (power) reste aussi conservée, puisque tous les patients sont inclus dans l'analyse statistique.
- Le risque de biais et de résultats faussement positifs des études est réduit (erreur dite de type 1). Nous allons expliquer cela un plus en détail:

Les violations du protocole par les patients ne sont souvent pas tout à fait le fruit du hasard. Elles sont souvent liées à la nature du traitement et au succès de ce dernier. On peut dire d'une façon générale que les patients qui suivent un protocole et qui terminent normalement une étude présentent de meilleurs résultats thérapeutiques que ceux qui le font pas. C'est aussi vrai pour les patients du groupe placebo des essais cliniques [5].

Les causes d'une mauvaise observance ou d'une interruption du traitement peuvent d'autre part être multiples, par exemple: a) une nette amélioration des symptômes et le fait de ne pas voir de justification à la poursuite du traitement, b) l'arrêt d'un traitement en apparence inefficace ou c) l'arrêt du traitement en raison d'effets indésirables intolérables. Comme les motifs ci-dessus ne se distribuent en règle générale pas de façon identique dans les différents groupes d'un essai, dans la mesure où ils dépendent directement de l'efficacité et de la tolérance des thérapies étudiées, une analyse des données excluant les sujets ayant interrompu le traitement impliquerait un biais – dans de nombreux cas à l'avantage du traitement testé.

L'exclusion de patients ne recevant pas le traitement prévu après la randomisation pose également un problème. Un exemple à ce propos: une étude compare un traitement chirurgical à un traitement conservateur médicamenteux. On présume que les deux traitements sont rigoureusement équivalents. L'exclusion des patients décédés après la randomisation, mais avant l'intervention chirurgicale, procurerait un avantage à la chirurgie et un procédé thérapeutique en réalité pas plus efficace pourrait ainsi être favorisé à tort. L'efficacité en apparence supérieure ne tiendrait en effet pas à une supériorité réelle, mais constituerait dans ce cas une conséquence de l'élimination systématique dans le groupe chirurgical des individus dont le pronostic est le plus mauvais [6].

Une analyse ITT correspond le mieux aux conditions réelles de la clinique

L'analyse ITT modifiée constitue une exception à cette règle. Le principe généralement admis est l'exclusion après coup des patients «randomisés à tort»¹.

- Une analyse ITT correspond le mieux aux conditions réelles de la clinique. Elle représente une approche conservatrice de l'analyse des résultats, dans le but d'éviter une surestimation de l'efficacité [7]. Dans la pratique quotidienne, les patients prennent aussi leurs médicaments de manière irrégulière ou les arrêtent prématurément. L'analyse ITT favorise donc la généralisation des résultats des essais cliniques. La comparaison des résultats des analyses ITT, per-protocole et as-treated des résultats d'un même essai clinique (pour autant que les trois méthodes soient applicables) confirme que l'ITT est l'approche la plus conservatrice et conduit habituellement aux différences les plus faibles entre les groupes de traitement (tab. 1 ◀; d'après [8]). Les analyses per-protocole et as-treated tendent au contraire à surestimer les effets thérapeutiques.

Limitations de l'approche ITT

Le principe de l'ITT est une stratégie visant à traiter les différents points évoqués ci-dessus de manière strictement logique. L'ITT ne reflète cependant pas non plus un état du «monde parfait». Voici quelques critiques à l'endroit de l'ITT et limitations de la méthode, ainsi que quelques exemples de situations qui justifient le recours à des méthodes d'analyse complémentaires ou posent l'indication à des approches carrément différentes.

- L'effet «réel» du traitement est tendanciellement sous-estimé par l'analyse ITT. Pour pouvoir évaluer cet effet ou l'effet «purement» biologique d'un traitement, on complètera l'évaluation par une analyse per-protocole.
- Dans les essais de bioéquivalence ou de non-infériorité, l'analyse per-protocole constitue l'approche la plus conservatrice. Une analyse ITT est plutôt en mesure d'effacer certaines différences entre les traite-

1 L'exclusion post-hoc de patients «randomisés à tort» de l'analyse statistique n'est acceptable que s'il ne peut pas être établi avec certitude au moment de la randomisation que tous les critères d'inclusion sont vérifiés. On peut citer pour exemple des essais comparatifs d'antibiotiques chez des patients atteints d'infections sévères nécessitant l'introduction immédiate d'un traitement – avant même que les résultats de l'hémoculture soient disponibles. Dans le cadre d'une étude chez des patients atteints d'infections cutanées, on peut ainsi supposer l'existence d'une infection à Gram⁺ (critère d'inclusion) et démarrer une antibiothérapie appropriée. Il est probable qu'on trouvera par la suite chez un petit nombre de patients une infection à Gram⁻ ou une infection mixte. Ces patients peuvent en principe être exclus de l'analyse finale des résultats. On parle dans ce cas d'une analyse ITT modifiée, ne portant que sur les patients ayant réellement répondu aux critères d'inclusions fixés préalablement. Il est néanmoins utile de procéder en plus à une analyse ITT incluant tous les patients, dans la mesure où cette situation est plus proche de la réalité clinique. Les autorités de contrôle des médicaments exigent d'ailleurs dans ces cas de figure des analyses ITT complémentaires tenant compte de tous les patients randomisés.

ments, d'augmenter le risque d'erreur de premier type et donc de susciter des conclusions fausses en favorisant l'équivalence [9]. Il y a beaucoup d'arguments en faveur d'une analyse per-protocole complémentaire. Les résultats et l'interprétation d'une étude d'équivalence ou de non-infériorité sont d'autant plus crédibles que les conclusions des analyses ITT et per-protocole sont proches [10].

- Pour évaluer la fréquence et le degré de sévérité des effets indésirables, on privilégie la méthode as-treated.
- Une analyse as-treated complémentaire se justifie lorsque le taux de cross-over des traitements est élevé.
- Le principal problème lié aux analyses ITT tient toutefois au fait que l'ITT n'est fondamentalement pas à même de minimiser le biais inhérent aux données manquantes lors des end-points des études. Si le taux loss-to-follow-up est important, autrement dit s'il y a beaucoup de données manquantes à l'issue d'une étude, la valeur de cette dernière est fortement compromise [4]. Le problème des données incomplètes ne peut être adressé que par des méthodes elles-mêmes sujettes à des limitations et impliquant en principe des biais significatifs (cf. paragraphe suivant). Cet aspect souligne une fois de plus l'importance que revêt une collecte de données aussi complète que possible.

Traitement des données manquantes

Il existe plusieurs méthodes statistiques pour traiter le problème des données manquantes. On distingue fondamentalement les stratégies d'imputation numériques (par ex. la méthode «Last Observation Carried Forward» [LOCF]) et les méthodes fondées sur les rangs (par ex. prise en compte des cas extrêmes, telles que l'analyse «best-and-worst-case»). Nous nous bornons ici à aborder brièvement quelques exemples.

La méthode LOCF consiste à répéter jusqu'à la fin de l'étude la dernière valeur post-randomisation mesurée.

Cette méthode intègre donc l'information obtenue plus tôt durant l'étude. Elle part cependant de l'hypothèse tout à fait irréaliste selon laquelle les valeurs relatives à un patient resteraient constantes. Un exemple va montrer que la méthode LOCF peut par ailleurs souffrir d'un biais important et qu'elle augmente le risque de faux-positifs: dans une étude fictive, des patients randomisés

dans le groupe interventionnel sont éliminés principalement au début du protocole à la suite d'effets indésirables sévères; or, si s'agit dans le même temps des patients qui répondaient le mieux au traitement. Les sujets ayant interrompu l'essai dans le groupe placebo sont en revanche surtout des patients dont l'état clinique allait s'aggravant. Il est évident que l'efficacité du traitement testé sera nettement surestimée par la méthode LOCF. D'abord les

drop-outs dans les deux groupes ne sont pas comparables; ensuite, il n'est pas réaliste de penser que les bonnes valeurs des patients du groupe interventionnel vont se maintenir après l'arrêt du traitement.

Les méthodes fondées sur les rangs s'utilisent dans des essais au cours desquels il manque des informations relatives à des variables dites dichotomiques (par ex. survivant ou décédé). Souvent, on procède à des analyses best-and-worst-case (analyse de la sensibilité), autrement dit des évaluations d'effets sur la base des valeurs extrêmes possibles. Plus le taux de drop-outs avec issue incertaine dans le cadre de l'étude est élevé, plus grande est l'incertitude des résultats estimés de cette façon. Dans l'optique du worst-case, l'effet réel est sous-estimé dans une proportion indéterminée et dans l'optique du best-case c'est l'inverse. Si certains résultats deviennent neutres par l'approche worst-case ou s'ils passent même du côté opposé, on ne pourra en définitive pas savoir si le traitement est inefficace ou si le résultat observé est simplement la conséquence des interruptions d'étude. L'exemple présenté dans le tableau 2 (d'après [11]) illustre les difficultés inhérentes à cette approche.

L'avantage de toutes les méthodes utilisées pour traiter les données manquantes réside dans le fond uni-

L'avantage de toutes les méthodes utilisées pour traiter les données manquantes réside dans le fait qu'elles sont supérieures à une simple exclusion de certains patients de l'analyse statistique

Tableau 1. Comparaison entre la chirurgie de by-pass et le traitement médicamenteux chez des hommes avec angor stable; patients survivants ou décédés à 2 ans post-randomisation (d'après [8]).

Selon randomisation	Intervention attribuée et réellement appliquée				Différence de mortalité (médicaments vs chirurgie)
	Médicaments		Chirurgie		
Effectivement appliqué	médicaments (a)	chirurgie (b)	chirurgie (c)	médicaments (d)	
Nombre de survivants	296	48	353	20	
Nombre de décès	27	2	15	6	
Mortalité	8,4%	4,0%	4,1%	23,1%	
ITT (a + b) vs (c + d)	7,8% (29/373)	5,3% (21/394)			2,4%
Per-protocole (a) vs (c)	8,4% (27/323)	4,1% (15/368)			4,3%
As-treated (a + d) vs (b + c)	9,5% (33/349)	4,1% (17/418)			5,4%

La mortalité des patients est la plus élevée dans le groupe des patients randomisés dans le groupe chirurgical, mais qui n'ont pas pu être opérés et qui ont finalement reçu le traitement médicamenteux (23,1%). L'analyse ITT attribuant ces patients au groupe sous traitement chirurgical, démontre un avantage infime pour la chirurgie versus traitement conservateur (médicamenteux). Dans l'analyse per-protocole, ces patients ne sont pas pris en compte en raison de violations du protocole. L'analyse as-treated indique la différence la plus nette en faveur de la chirurgie, car elle attribue les patients mentionnés ci-dessus au groupe traité par médicaments.

Tableau 2. Comparaison entre la dilatation au ballonnet et l'implantation de stents chez 110 patients avec occlusion du greffon vasculaire après by-pass coronarien (d'après [11]).

	Angioplastie (n = 110)	Stent (n = 110)	Différence entre les taux de resténose (du point de vue du traitement par stent) et intervalle de confiance à 95%	
Resténoses à 6 mois				
Oui	37	32		
Non	43	54		
Inconnu	30	24		
Taux de resténoses en %				
Patients avec status connu	46% (37/80)	37% (32/86)	-9%	(-24 à + 6)
Analyse selon le best-case pour les stents	61% (67/110)	29% (32/110)	-32%	(-44 à -19)
Analyse selon le worst-case pour les stents	34% (37/110)	51% (56/110)	+17%	(+ 4 à + 30)

Chez 30 patients, c'est-à-dire chez près d'un tiers des patients, le status vasculaire à 6 mois était inconnu. La signification d'une analyse de patients dont l'état des coronaires est connu, ainsi que les analyses de type best- and worst-case est par conséquent limitée. Chez les patients avec état coronarien documenté, le traitement par stent entraîne une réduction de 9% des resténoses par rapport à la dilatation par ballonnet seule. Dans une analyse complémentaire de type worst-case, cette différence s'inverse en faveur de l'angioplastie. Il est donc pratiquement impossible d'évaluer le «véritable» effet.

quement dans le fait qu'elles sont malgré tout supérieures à une simple exclusion de certains patients de l'analyse statistique. Parmi les inconvénients, on citera le risque de biais et de faux-positifs ou de faux-négatifs, ainsi que le haut degré d'incertitude concernant le véritable effet et la perte de puissance des tests [3, 12].

Conclusions

Pour les médecins impliqués dans la planification et la réalisation d'essais cliniques:

Le screening des patients avant la randomisation est extrêmement important. L'étude sera conçue de manière à assurer une forte adhérence au protocole et à éviter autant que possible les données manquantes – par exemple en évitant des visites trop fréquentes ou des investigations diagnostiques inutiles. S'agissant de l'end-point de l'étude, il faut essayer de suivre les patients de manière optimale jusqu'à la fin planifiée du protocole ou jusqu'à la survenue d'un événement fatal prématuré – même lorsque la poursuite du traitement n'est pas possible, par ex. à la suite de manifestations

d'intolérance. Le protocole d'étude devrait déjà mentionner dans quelle mesure on s'attend à des données manquantes et indiquer quelles méthodes statistiques seront utilisées pour les traiter [13].

Pour les lecteurs de publications scientifiques:

Pour pouvoir estimer la vraie valeur d'une étude, il faut poser les bonnes questions. Une analyse ITT a-t-elle été appliquée à tous les patients randomisés? Quel est le taux d'abandons? Comment a-t-on traité le problème des interruptions de l'étude et des données manquantes dans l'analyse statistique?

S'il manque des données à l'end-point pour une proportion significative des patients, on considérera avec une grande retenue l'affirmation qu'une analyse ITT a été utilisée et on se méfiera par conséquent des résultats publiés.

Correspondance:

Dr Peter Kleist
Glaxo Smith Kline AG
Talstrasse 3-5
CH-3053 Münchenbuchsee
peter.m.kleist@gsk.ch

Références recommandées

- Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134:663-94.
- Fergusson D, Aaron SD, Gyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ.* 2002;325:652-4.

- Gotzsche PC. Lessons from and cautions about noninferiority and equivalence randomised trials. *JAMA.* 2006;295:1172-4.

Vous trouverez la liste complète et numérotée des références dans la version en ligne de cet article sous www.medicalforum.ch.

Références 11-09 (online version)

- 1 Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134:663–94.
- 2 International Conference on Harmonisation (ICH). Guideline E9: Statistical Principles for Clinical Trials. www.ich.org/LOB/media/MEDIA485.pdf.
- 3 Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ.* 1999;319:670–4.
- 4 Montori VM, Guyatt GH. Intention-to-treat principle. *CMAJ.* 2001;165:1339–41.
- 5 Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *N Engl J Med.* 1980;303:1038–41.
- 6 Bhandari M, Gyatt GH, Swiontkowski MF. User's guide to the orthopaedic literature: how to use an article about a surgical therapy. *J Bone Joint Surg Am.* 2001;83:916–26.
- 7 Fergusson D, Aaron SD, Gyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ.* 2002;325:652–4.
- 8 European Coronary Surgery Study Group. Coronary-artery bypass surgery in stable angina pectoris: Survival at two years. *Lancet.* 1979;i:889–93.
- 9 Gotzsche PC. Lessons from and cautions about noninferiority and equivalence randomised trials. *JAMA.* 2006;295:1172–4.
- 10 Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ.* 1996;313:36–9.
- 11 Savage MP, Douglas JS Jr, Fischman DL, et al. Stent placement compared with balloon angioplasty for obstructed coronary bypass grafts. *N Engl J Med.* 1997;337:740–7.
- 12 Streiner D, Geddes J. Intention to treat analysis in clinical trials when there are missing data. *Evid based Ment Health.* 2001;4:70–1.
- 13 Committee for Proprietary Medicinal Products (CPMP). Points to consider on missing data. London, 15 November 2001. www.emea.europa.eu/pdfs/human/ewp/177699EN.pdf.