

# Les différentes méthodes d'évaluation des nouvelles technologies et des dispositifs médicaux

## Different Methods of Assessment of New Technologies and Medical Devices

Alain Bernard

*Service de chirurgie thoracique - CHU Bocage Central - 14, rue Gaffarel - BP 77908 - Dijon Cedex.*

### Mots clés

- ◆ Essais contrôlés randomisés
- ◆ Évaluation
- ◆ Dispositifs médicaux

### Résumé

Les données cliniques fournies pour l'évaluation des dispositifs médicaux (DM) sont habituellement insuffisantes. Tout nouveau DM devrait faire l'objet d'études cliniques de qualité pour la démonstration de son intérêt pour les patients. L'essai contrôlé randomisé est le plan expérimental offrant le meilleur niveau de preuve pour démontrer l'efficacité d'un nouveau DM par rapport aux thérapeutiques de référence. Ce type d'étude permet de réduire les risques de biais ne permettant pas de savoir si la différence observée est due à la nouvelle thérapeutique ou bien à d'autres facteurs non contrôlés. Cependant, le contexte clinique de certains DM rend difficile la réalisation d'un essai contrôlé randomisé conventionnel. D'autres plans expérimentaux peuvent être proposés pour la démonstration de l'efficacité.

Le plan de « Zelen » facilite le recrutement des patients mais il peut poser des problèmes éthiques. Les essais basés sur l'expertise favorisent l'acceptabilité des médecins mais la mise en place est plus contraignante. Les essais en « cross over » permettent de réduire le nombre de patients mais ils ne sont pas applicables lorsqu'une courbe d'apprentissage est nécessaire. Les essais séquentiels présentent l'avantage d'arrêter plus précocement un essai grâce à l'analyse intermédiaire cependant un comité de surveillance est nécessaire. Les essais bayésiens combinent les informations existantes et celles fournies par l'essai en cours. Ces méthodes sont intéressantes pour les petites populations. L'inconvénient est la prise en compte d'une information à priori erronée.

L'essai contrôlé randomisé est applicable dans un grand nombre de situations. Les méthodologies proposées permettent de répondre à la diversité des DM.

### Keywords

- ◆ Randomised controlled trial
- ◆ Assessment
- ◆ Medical device

### Abstract

Clinical evidence available for the assessment of medical devices (MD) is frequently insufficient. New MDs should be subjected to high quality clinical studies to demonstrate their interest for patients. The randomized controlled trial (RCT) is the study design reaching the highest level of evidence in order to demonstrate the efficacy of a new MD. This type of trial reduces the risk of bias that it does not reveal whether the observed differences are related to the new treatment or other factors. However, the clinical context of some MDs makes it difficult to carry out a conventional RCT. Other experimental designs can thus be proposed to demonstrate the clinical efficacy of new MDs.

Zelen's design trials and randomized consent design trials facilitate the recruitment of patients, but can rise ethical problems. Expertise-based randomized controlled trials involve randomizing to a team that specializes in a given intervention. Sometimes, the feasibility of an expertise-based randomized trial may be greater than that of a conventional trial. Crossover trials reduce the number of patients, but are not applicable when a learning curve is required. Sequential trials have the advantage of allowing to stop early a trial depending on the results of first inclusions, but they require an independent committee. Bayesian methods combine existing information with information from the ongoing trial. These methods are particularly useful in situations where the number of subjects is small. The disadvantage is the risk of including erroneous prior information.

Although, the RCTs are applicable in many situations. The previous methods can meet the diversity of situations existing with MDs.

Les données cliniques fournies pour l'évaluation des Dispositifs médicaux (DM) sont habituellement insuffisantes ou limitées (1). Ce constat est partagé au niveau Européen comme le montre de récentes publications (2,3). Tout nouveau DM devrait faire l'objet d'études cliniques de qualité pour la démonstration de son intérêt pour les patients. Les DM sont des produits de santé au même titre que les médicaments, destinés à traiter des patients. Enfin les performances techniques

d'un nouveau DM ne présagent en rien de son efficacité clinique pour le patient.

Le développement clinique de tout nouveau DM débute par les études de faisabilité qui sont proposées immédiatement à la suite de la phase préclinique (4). Ces études de faisabilité sont utiles pour préciser les caractéristiques des patients, pour la mise au point de la technique d'implantation et de déterminer le critère de jugement le plus approprié pour me-

### Correspondance :

*Pr Alain Bernard*

*Head of Department of Thoracic Surgery - Bocage Central - 14 rue Gaffarel - BP 77908 - 21079 Dijon Cedex France.*

*Tel: 33 3 80 29 33 52. Fax: 33 3 80 29 33 85 - E-mail : alain.bernard@chu-dijon.fr*

surer l'efficacité du DM. Ces étapes fournissent des informations essentielles pour le futur essai contrôlé randomisé pour la démonstration de l'efficacité.

L'essai contrôlé randomisé est le plan expérimental offrant le meilleur niveau de preuve pour démontrer l'efficacité d'un nouveau DM par rapport à une thérapeutique de référence (5). Dans la plupart des cas, la réalisation d'un essai contrôlé randomisé est possible pour le développement clinique d'un nouveau DM. Cependant, le contexte clinique de certains DM peut rendre difficile la réalisation d'un essai contrôlé randomisé en groupes parallèles conventionnel. Pour pallier ces difficultés, d'autres plans expérimentaux pourraient être proposés pour la démonstration de l'efficacité. Les objectifs de ce travail sont d'identifier ces différentes méthodes adaptées au développement clinique des DM, de préciser leurs limites et de proposer leur application.

## Les limites des essais contrôlés randomisés conventionnels

La réalisation d'un essai contrôlé randomisé en groupes parallèles conventionnel peut s'avérer difficile à mettre en place pour les raisons suivantes :

- Le choix du moment de l'évaluation peut influencer les résultats de l'évaluation. Un essai réalisé très précocement peut être influencé par les évolutions technologiques d'un nouveau DM. Dans ce cas, les résultats négatifs peuvent davantage refléter une technique non maîtrisée qu'une réelle inefficacité clinique (6). A contrario, lorsqu'une technologie de santé ou un DM sont largement diffusés, il devient difficile de faire adhérer les médecins ou les patients à un essai clinique (7). Cette situation est responsable de l'utilisation du DM ou de technologie de santé sans aucune preuve de son efficacité (7).
- La faible taille de la population éligible constitue une spécificité des études portant sur les DM (4,7). Le manque de puissance peut être responsable d'un résultat non significatif de l'essai pour la démonstration de l'efficacité thérapeutique ou la mise en évidence d'événements indésirables.
- L'acceptabilité des patients joue un rôle essentiel pour la réussite de l'étude. Les patients peuvent refuser de participer à un essai contrôlé randomisé lorsque celui-ci compare une technique chirurgicale invasive à une méthode non-invasive. La large diffusion d'un nouveau DM ou d'une nouvelle technologie peut avoir un impact sur les patients qui refuseront de participer à cet essai pensant bien souvent à tort, qu'il est supérieur au traitement de référence (6).
- L'aveugle est un élément important des essais cliniques afin de limiter un certain nombre de biais liés à la subjectivité du médecin ou du patient. Le plus souvent au cours des essais non pharmacologiques (8) l'aveugle est difficile à obtenir pour des raisons pratiques ; par exemples, la cicatrice révélant le type de chirurgie ou la perception de l'action du DM par le patient.
- Le comparateur doit être en priorité le traitement de référence établi à partir des preuves scientifiques (5). De manière exceptionnelle, l'absence de traitement efficace peut nécessiter d'avoir recours à un placebo (9). Dans ce cas de figure plusieurs options ont été décrits par Boutron I et al. (10) par exemple : la cicatrice peut être masquée par des pansements identiques ou l'utilisation du DM inactif.
- La courbe d'apprentissage peut jouer un rôle important sur les performances cliniques de certaines technologies de santé ou DM (7). Un manque d'expérience peut influencer négativement le résultat de l'essai (11). La courbe d'apprentissage devra être prise en compte au moment de la rédaction du protocole. Plusieurs solutions sont possibles :
  - de ne retenir que des centres ayant un certain niveau d'expérience ;
  - de standardiser la technique afin de limiter la variabilité

entre les centres ;

- la prise en compte dans le plan expérimental.

## Les différentes méthodes

L'évaluation en insu du critère de jugement, plusieurs méthodes peuvent être proposées ; la personne qui recueille le critère de jugement est indépendante de l'équipe qui a utilisé le DM ou réalisé l'acte technique ; l'évaluateur ne devrait pas pouvoir identifier le traitement attribué au patient (8). La mesure peut être centralisée pour des examens biologiques ou radiologiques ou des extraits de l'examen clinique (photos de plaie par exemple). Un dernier moyen pour compenser l'absence d'aveugle est l'utilisation d'un comité d'adjudication indépendant des investigateurs pour valider le critère de jugement. Cette dernière méthode qui est efficace en l'absence d'aveugle, alourdit la réalisation de l'étude.

### Le plan de « Zelen » (tableau I)

Seuls les patients randomisés dans le groupe du nouveau traitement devront signer le consentement éclairé. Si le patient refuse, il recevra le traitement de référence (12). Trois variantes du plan de « Zelen » ont été proposées selon le contexte clinique (14). L'avantage de ce plan expérimental est de faciliter l'inclusion des patients en limitant leur refus (15). Le recrutement des patients est facilité même en cas de forte préférence pour un des traitements. Ce type d'étude n'altère pas la confiance que le patient peut avoir en son médecin (16). Plusieurs inconvénients ont été rapportés concernant ce type de plan expérimental. Le premier est éthique puisque seuls les patients ayant eu le nouveau traitement, signeront le consentement éclairé. Ce type de plan expérimental peut diluer l'effet du nouveau traitement avec le risque de ne pas pouvoir conclure si les patients randomisés refusent le nouveau traitement au profit de la thérapeutique de référence (12). Un biais de sélection est possible avec le risque d'une sous-représentation des patients à mauvais pronostic dans le groupe expérimental. Les patients à risque élevé peuvent refuser le nouveau traitement au profit de la thérapeutique de référence (12). Actuellement peu d'essais de ce type ont été proposés en France pour des raisons éthiques. Le champ d'application possible concerne les essais où il n'y a pas de visites supplémentaires pour le traitement de référence et lorsque le décès est le critère de jugement principal (16). Il est adapté pour les essais en chirurgie et les DM (17). Un essai comparant la radiofréquence à la chirurgie pour le traitement des petits hépato-carcinomes, a utilisé un plan de Zelen (17). Les auteurs justifient l'utilisation de ce plan expérimental pour prévenir le refus des patients d'être traités par la radiofréquence alors que la chirurgie est la référence (17). D'après le flow chart, quatre patients refuseront de participer à l'essai (17). Après la randomisation le nombre de patients refusant la thérapeutique proposée est bien équilibré, quatre dans le groupe radiofréquence et quatre dans le groupe chirurgie (17). L'analyse en intention de traiter à porter sur 168 patients (84 dans chacun des groupes). Les caractéristiques des patients étaient comparables dans les deux groupes thérapeutiques.

### Essais basés sur l'expertise (tableau I)

Le principe est de randomiser les patients auprès du chirurgien ou de l'équipe qui maîtrise l'intervention ou la technique (18). Dans ce cas le médecin ne réalise que la procédure qu'il maîtrise bien. Ce type d'étude nécessite la présence d'un expert de chacune des interventions dans le même centre (19). De plus la consultation initiale de pré-inclusion devra être réalisée par une personne neutre pour déterminer l'éligi-

Plan	Principes	Avantages	Inconvénients
Plan de « Zelen »	Randomisation avant le consentement du patient	<ul style="list-style-type: none"> <li>• Faciliter l'inclusion</li> <li>• N'altère pas la confiance patient/ médecin</li> </ul>	<ul style="list-style-type: none"> <li>• Biais de sélection possible</li> <li>• Problèmes éthiques</li> </ul>
Essais basés sur l'expertise	Randomiser les patients auprès des praticiens qui maîtrisent la technique	<ul style="list-style-type: none"> <li>• Meilleure acceptabilité</li> <li>• Limiter les biais d'exécution</li> <li>• Réduire les écarts au protocole</li> </ul>	Difficulté de savoir si la supériorité est le fait de la technique ou de l'expertise du praticien
« Tracker trial design »	Révision du protocole au cours de l'essai	Réaliser une évaluation très précoce	Difficile à mettre en pratique pour l'analyse et la logistique
Essai en cluster	Randomiser un groupe d'individus (hôpital, service)	Plus simple à mettre en œuvre	<ul style="list-style-type: none"> <li>• Manque de puissance</li> <li>• Recrutement différent entre les clusters (biais de sélection)</li> </ul>
Essai en cross-over	La randomisation concerne la séquence d'administration du traitement	Réduire le nombre de sujets nécessaires	<ul style="list-style-type: none"> <li>• Le traitement en premier a un effet rémanent au cours de la deuxième période</li> <li>• Pathologie stable</li> <li>• Le critère de jugement ne peut pas être la guérison ou le décès</li> </ul>
Les essais séquentiels	Les résultats de l'essai concernant les patients inclus (analyse intermédiaire) sont analysés avant la randomisation de nouveaux patients	Minimisation du nombre de sujets nécessaires	<ul style="list-style-type: none"> <li>• Un seul critère de jugement</li> <li>• Manque de puissance pour les critères de jugement secondaires notamment les événements indésirables ou complications</li> <li>• Le délai entre l'inclusion et la mesure du critère de jugement doit être court</li> <li>• Nécessité d'un comité de surveillance</li> </ul>
Essai adaptatifs	<ul style="list-style-type: none"> <li>• Le schéma de l'étude peut être changé sur la base des résultats observés au cours de l'analyse intermédiaire</li> <li>• Les adaptations peuvent porter sur le nombre de patients, le critère de jugement, l'ajout ou la suppression d'analyse intermédiaire</li> </ul>	<ul style="list-style-type: none"> <li>• Réduction du nombre de sujets nécessaires</li> <li>• Une plus grande flexibilité</li> </ul>	<ul style="list-style-type: none"> <li>• Contraintes logistiques plus grandes</li> <li>• Nécessité d'un comité de surveillance</li> <li>• Critère de jugement unique</li> <li>• Remise en cause de la validité interne</li> </ul>
Essais bayésiens	<ul style="list-style-type: none"> <li>• Combinaison des informations existantes (à priori) et les données fournies par l'essai en cours</li> <li>• Les informations à priori proviennent de la littérature ou avis d'experts modélisés</li> </ul>	<ul style="list-style-type: none"> <li>• Plus flexible</li> <li>• Réduire le nombre de sujets nécessaires</li> <li>• Utilisation de toute l'information disponible</li> </ul>	La prise en compte erronée de l'information existante (à priori)

Tableau I : Synthèse des différents plans expérimentaux.

bilité des patients (19). Ce plan requiert une stratification par chirurgien et par centre (19). Parmi les avantages décrits, cette méthode permettrait de limiter les biais d'exécution liés à l'absence d'aveugle (20). Le risque d'avoir un suivi différent (traitements associés ou prise en charge) dans les deux groupes est plus limité qu'au cours d'un essai conventionnel. Ce plan expérimental pourrait réduire les écarts au protocole (19). Ils améliorent la faisabilité, les médecins n'ont pas besoin d'être formés aux deux procédures (19). Enfin dernier avantage et non des moindres concerne l'acceptabilité par les médecins qui seront moins réticents à participer à un essai (19,18). L'inconvénient de ce plan expérimental réside dans la difficulté de savoir si la différence observée est le fait de la technique ou de l'expertise du médecin (19). Ce type de plan expérimental est intéressant lorsqu'une courbe d'apprentissage est nécessaire pour maîtriser la technique et ce d'autant que le niveau d'expertise peut avoir une influence sur le résultat. Par exemple, un essai multicentrique compare l'arthroscopie à la chirurgie ouverte pour le traitement de la coiffe des rotateurs (21). Sur le même site, au chirurgien sachant réaliser l'intervention sous arthroscopie était associé un chirurgien pratiquant la voie ouverte. L'objectif étant de minimiser l'impact de l'expertise sur les résultats de l'essai.

### « Tracker trial design » (tableau I)

Cet essai a été développé afin de prendre en compte les évolutions technologiques des DM (22). Des modifications au cours de l'essai sont autorisées et prises en compte dans l'analyse statistique. Les variations dans le schéma d'étude (révision du protocole) sont permises (22). Le principal avantage est de réaliser une évaluation très précoce des nouveaux DM et des nouvelles technologiques avant une large diffusion

(22). Ce type d'essai possède un certain nombre d'inconvénients, les méthodes sont plus sophistiquées que pour les essais conventionnels, l'organisation pratique est difficile avec un impact sur le budget de l'étude (22). Actuellement il existe peu d'exemples disponibles dans la littérature sur ce type d'étude.

### Essais en « cluster » (tableau I)

Le principe est de randomiser des groupes d'individus (hôpital, service...). A chaque centre lui sera attribué un des deux traitements de manière aléatoire (23). Ce plan expérimental est facile à mettre en œuvre sur le plan logistique. Il peut avoir un intérêt pour la comparaison de stratégies. Il présente des inconvénients comme le manque de puissance (23), lorsque le recrutement est différent entre les clusters en termes de nombre et de caractéristiques de patients. Parfois certains clusters peuvent demeurer vides lorsque la personne responsable n'inclut pas de patients (24). Ces situations mettent en péril la comparabilité des groupes (biais de sélection) et remettent en cause la validité interne de l'essai (25). L'application de ce genre d'étude reste limitée pour le développement clinique du DM.

### Les essais en « cross over » (tableau I)

La randomisation concerne la séquence d'administration du traitement. Chaque patient reçoit ainsi les deux traitements (26). Ce plan expérimental a l'avantage de réduire le nombre de sujets nécessaires (26). La préoccupation majeure est le risque d'effet « carry-over » qui est du à l'effet rémanent du traitement administré en premier au cours de la deuxième période. Ces essais ne sont pas possibles lorsque le critère de

jugement est la guérison ou le décès (26) ou lorsqu'une courbe d'apprentissage est nécessaire. Pour utiliser un essai en cross over il est nécessaire que la pathologie soit stable. Ce type d'essai peut avoir un intérêt en dermatologie, cosmétologie et pour certains DM implantables actifs comme par exemple les neuro-stimulateurs.

### Les essais séquentiels (tableau I)

Ils ont été développés pour arrêter des essais plus rapidement. Ils font partie des techniques d'analyse intermédiaire des essais contrôlés randomisés. Les résultats de l'essai concernant les patients inclus sont analysés avant la randomisation de nouveaux patients. Plusieurs méthodes ont été décrites (27-31). Ces méthodes présentent l'avantage d'arrêter plus précocement un essai (32). Le nombre moyen de sujets nécessaires est plus faible que dans les essais conventionnels (33,34). Ils possèdent un certain nombre de contraintes : un seul critère de jugement, le délai entre l'inclusion des patients et la mesure du critère de jugement doit être court. Un arrêt précoce risque d'engendrer un manque de puissance pour certains critères de jugement secondaires. Un comité de surveillance indépendant devrait être mis en place pour décider la poursuite ou l'arrêt de l'essai (33). Il est conseillé d'avoir des inclusions régulières avec un suivi de bonne qualité pour éviter des délais dans les mises à jour des données. Ces techniques ont un intérêt particulier lorsque l'on s'intéresse à des populations cibles restreintes.

### Les essais avec randomisation adaptative (tableau I)

L'objectif est d'optimiser le nombre total de patients traités avec la meilleure des thérapeutiques. Au début de l'essai les patients sont randomisés avec un ratio 1 : un dans chacun des bras puis ensuite lorsqu'une différence commence à être observée, le ratio est modifié en faveur du groupe qui reçoit le traitement semblant être plus efficace (32). Ces essais reposent sur la réalisation d'analyses intermédiaires. Le plan expérimental peut être changé sur la base des résultats obtenus lors des analyses intermédiaires (35). Les adaptations possibles portent sur la réévaluation du nombre de sujets nécessaires, sur le critère de jugement, l'ajout ou la suppression d'analyses intermédiaires (33). Les essais adaptatifs peuvent être construits en deux phases ou en plusieurs phases (36). Au cours des essais en deux phases, dans un premier temps un échantillon de taille N1 est sélectionné. Selon les résultats, soit l'étude est interrompue soit l'étude est poursuivie. Il est nécessaire de constituer un comité de surveillance indépendant qui décidera des adaptations du nombre de sujets nécessaires et de la poursuite ou de l'arrêt de l'inclusion (33,36). Un essai adaptatif requiert des contraintes logistiques plus grandes qu'un autre type d'essai (36). Le critère de jugement devra être simple et unique (32). Les essais adaptatifs permettent une plus grande flexibilité et une diminution du nombre de sujets nécessaires. Cependant la validité interne de ces essais est remise en question, il pourrait avoir les mêmes inconvénients que les témoins historiques (27,32). Ils ne seraient jugés pas plus intéressants que les essais séquentiels selon certains auteurs (32,36,37). Ce type d'essai pourrait être intéressant dans le cadre de technique nécessitant une courbe d'apprentissage ou lorsqu'un nouveau DM subit des évolutions technologiques.

### Les essais bayésiens (tableau I)

Au cours d'un essai contrôlé randomisé conventionnel, les données issues des études antérieures sont utilisées pour la conception de l'essai. Au moment de l'analyse, elles ne sont pas utilisées, seules les données de l'essai en cours sont prises

en compte. Au contraire, l'approche bayésienne combine des informations existantes et les données fournies par l'essai en cours (38). Elle utilise des informations déjà existantes, en priorité la littérature, en l'absence de celle-ci un avis d'experts modélisé. A partir du théorème de Bayes, les probabilités à posteriori sont estimées à partir des données de l'essai en cours, conditionnellement aux informations antérieures (à priori) (38). L'estimation bayésienne ne fournit pas d'intervalle de confiance mais un intervalle de crédibilité (38). A la grande différence de l'approche dite « fréquentiste », il n'y a pas de tests statistiques mais des résultats avec un intervalle de crédibilité. Ces méthodes offrent une certaine souplesse. Elles peuvent être utilisées au cours des essais adaptatifs en réalisant des analyses intermédiaires (39). L'utilisation de ces méthodes bayésiennes est préconisée pour l'évaluation clinique des DM (40,41). L'avantage de ces méthodes permet de réduire le nombre de sujets nécessaires particulièrement intéressant pour les petites populations cibles. L'inconvénient majeur est la prise en compte arbitraire pouvant être erronée d'une information à priori qui influencera inévitablement le résultat final. Ce genre de situation reviendrait à inclure dans une méta-analyse une importante étude biaisée en faveur de l'effet traitement, en compensant totalement les résultats des autres études non biaisées. Un exemple intéressant a été publié en 2010, un essai contrôlé randomisé bayésien comparant les médicaments anti-arythmiques à la radiofréquence pour le traitement des fibrillations atriales paroxystiques (42). Le nombre maximum de sujets nécessaires était estimé à 230 patients. Des analyses bayésiennes intermédiaires ont été programmées pour 150 patients, puis 175 et 200 patients. L'essai a été arrêté à 150 patients car l'analyse intermédiaire bayésienne a montré une probabilité prédictive de succès grâce à la radiofréquence de plus 99.9 % au-delà du seuil fixé par le protocole.

### Les études observationnelles comparatives non randomisées

Dans le cadre du développement clinique de tout nouveau DM, ce type d'étude doit demeurer l'exception et ne les réserver qu'aux cas où un essai contrôlé randomisé est impossible à réaliser.

#### « Comprehensive cohort design »

Tous les patients éligibles pour un essai sont recrutés indépendamment de leur consentement à la randomisation (43). Pour les patients qui refusent la randomisation, ils reçoivent le traitement qu'ils préfèrent. Pour ce type d'étude la variable « randomisation » est introduite dans l'analyse comme une variable d'ajustement. Elle permet d'améliorer le recrutement des patients, de favoriser l'acceptabilité et de réduire le risque de sélectionner des patients peu représentatifs (44). Ce plan expérimental ne garantit pas que les patients soient comparables dans les deux groupes (biais de sélection) comme dans les études observationnelles.

### Etude observationnelle comparative prospective

Le principe est de comparer au sein d'une cohorte de patient, les patients ayant eu le traitement A ou le traitement B. La répartition des patients dans chacun des groupes n'est pas établie par l'investigateur. C'est simplement une donnée observée. Ce type d'étude ne garantit pas la comparabilité des deux groupes en particulier vis-à-vis des variables non mesurées (biais de confusion). La différence observée entre les deux groupes pourrait être due à d'autres facteurs que l'effet du traitement. Ce type d'étude ne permet pas de conclure en termes de causalité sur la supériorité d'une thérapeutique par rapport à une autre. La seule conclusion auto-

	Évolution technologique	Faible population	Événements indésirables	Évolution technologique	Faible population
Plan de « Zelen »					/
Essai basé sur l'expertise				/	
Essais en « cross over »		/			
Essais en « cluster »				/	
Essais adaptatifs	/	/			
Essais séquentiels		/	/		
« Comprehensive cohort design »					/
Essais bayésiens	/	/			

Tableau II : Les plans expérimentaux adaptés à la problématique du DM.

risée, c'est d'observer qu'il existe une différence entre les deux traitements. Pour atténuer le risque de biais de confusion, le score de propension a été développé (45). Le score de propension correspond à la probabilité conditionnelle pour un patient de recevoir le traitement A plutôt que le traitement B. L'objectif est d'équilibrer la distribution des covariables (âge, sexe, comorbidités,...) entre les deux groupes afin de neutraliser les facteurs de confusion (46,47). La comparaison entre les deux groupes se fait au sein de la même classe du score de propension (45). Une fois établi le score de propension peut être utilisé pour réaliser un appariement, une stratification, un ajustement ou une pondération (48). Le score de propension ne peut être réalisé que sur des variables recueillies (49) avec le risque d'obtenir des groupes non comparables (49). Une analyse de sensibilité permet d'identifier d'éventuels biais cachés (50). Une revue systématique comparant les résultats des essais contrôlés randomisés à ceux des études observationnelles utilisant un score de propension a été publiée sur les différentes thérapeutiques au cours des syndromes coronariens aigus (51). Globalement les études observationnelles rapportent une plus grande quantité d'effet que les essais contrôlés randomisés pour la mortalité à long terme. Il est à noter qu'un seul essai contrôlé randomisé parmi les 10 rapportait des résultats significatifs. A l'inverse neuf études observationnelles sur 10 étaient statistiquement significatives.

## Discussion

La revue des différentes méthodologies existantes montre que les conditions sont favorables pour pouvoir réaliser un développement clinique pour tout nouveau DM. Tout progrès médical ne peut se concevoir qu'en se comparant aux thérapeutiques existantes. C'est la seule façon de confirmer que ce nouveau DM permet d'améliorer la santé des patients par rapport aux traitements conventionnels et ne pas maintenir la soit disant nouveauté dans le domaine des hypothèses. La réalisation d'un essai comparatif devrait devenir la norme pour tout nouveau DM à risques. Bien évidemment, un industriel a tout intérêt de réaliser un essai comparatif s'il estime que son DM constitue un progrès médical pour les patients. La réalisation d'un essai de qualité est une contrainte mais au bout du compte c'est un investissement pour l'avenir et l'assurance de valoriser le DM en montrant qu'il permet de mieux traiter les patients. Contrairement à ce que l'on peut lire, un essai comparatif ne constitue pas un frein à l'innovation. Les différents plans expérimentaux proposés sont adaptés à une faible population cible, aux évolutions technologiques de certains DM et à la nécessité d'une courbe d'apprentissage. Selon les circonstances, un ou plusieurs plans expérimentaux peuvent être proposés afin d'aider les porteurs de projet et les industriels. Elles dépendent des caractéristiques du dispositif médical, de l'acceptabilité médicale et de l'acceptabilité du patient :

## Les caractéristiques du dispositif médical

- Lorsque le DM doit subir des évolutions technologiques importantes et rapides, au cours d'un développement clinique très précoce, plusieurs plans expérimentaux peuvent être proposés (tableau II) : les essais avec randomisation adaptative à condition que le critère de jugement soit unique et les essais bayésiens. Lorsque la population cible est faible, les essais bayésiens sont adaptés ainsi que les essais en cross-over également à la condition que la pathologie étudiée soit stable et que le critère de jugement puisse se répéter et enfin les essais séquentiels peuvent être envisagés. Lorsqu'un DM est potentiellement porteur d'événements indésirables, les essais séquentiels permettraient d'arrêter suffisamment précocement l'étude.
- L'acceptabilité médicale est essentielle pour la réussite d'un essai et notamment lorsque la maîtrise de la technique influence le résultat ou lorsque les techniques sont très différentes ou lorsqu'une courbe d'apprentissage est nécessaire. L'essai basé sur l'expertise est tout à fait adapté à cette problématique (tableau II). Exceptionnellement les essais en cluster pourraient être utilisés.
- L'acceptabilité du patient peut se poser lorsque le comparateur est une technique invasive ou la technique est déjà largement diffusée (tableau II). Le plan de « Zelen » permet de favoriser l'acceptabilité des patients. En dernier recours l'essai basé sur les préférences des patients « comprehensive cohort design » pourrait être utilisé tout en sachant qu'il ne garantit pas la comparabilité des deux groupes de patients.
- A la suite de cette revue des différents plans expérimentaux, la place des études observationnelles est très réduite dans le cadre d'un développement clinique pour un nouveau DM dit à risque. Les études observationnelles même avec une méthodologie rigoureuse ne peuvent pas garantir l'absence de biais. Les interprétations de ce type sont plus limitées sur la démonstration de la supériorité du nouveau DM par rapport à la thérapeutique de référence. Les études observationnelles pourraient être proposées lorsque la technique ou le DM sont largement diffusés où l'acceptabilité de tout plan expérimental avec randomisation s'avérerait un échec. Il vaut mieux une étude observationnelle qu'aucune preuve scientifique pour la démonstration de l'intérêt clinique du DM.
- Le constat qui s'impose, toutes les méthodes existent pour réaliser une évaluation clinique d'un nouveau DM. Cependant, l'accompagnement et la formation des porteurs de projet ou industriel est indispensable. Ces méthodologies sont beaucoup moins connues que les essais conventionnels. La mise en place d'un essai est une contrainte en termes d'organisation et de coûts financiers. Mais il vaut mieux perdre un peu de temps au début afin de proposer un essai de qualité pour la démonstration de l'efficacité. Cet investissement permettra une valorisation du nouveau DM en démontrant le progrès médical qu'il constitue pour les pa-

tients. Et au bout du compte la société et la puissance publique auront moins de réticences pour le prendre en charge. C'est un changement de paradigme que l'on assistera en matière d'évaluation clinique de DM si les industriels ou porteurs de projet acceptent de réaliser un essai clinique de qualité rapidement dès que la mise au point industriel vient de s'achever.

## Conclusion

L'essai contrôlé randomisé conventionnel est applicable dans un grand nombre de situations. La grande diversité et l'hétérogénéité des situations font que ces essais conventionnels ne peuvent pas toujours être appliqués au développement clinique de certains DM dits à risque. Pour répondre à toutes ces difficultés, ce papier propose une boîte à outils afin de permettre à chacun de trouver une réponse à sa problématique.

## Discussion en séance

Question de M Laskar

Quelle serait la validité d'une étude si elle était transposée dans un autre pays ?

Réponse

Une étude réalisée dans un autre pays est recevable, la seule restriction concerne l'extrapolation des résultats aux patients français. Pour accepter l'extrapolation des résultats, il faut pouvoir répondre à deux questions importantes :

- les patients de l'étude sont-ils comparables aux patients français ?
- la pratique médicale de ce pays est-elle la même que celle de la France ?

Question de Fr Richard

Que pensez-vous de la prise en charge des nouvelles technologies ?

Réponse

Les nouvelles technologies ne devraient être utilisées qu'après avoir été évaluées ou dans le cadre d'essai clinique. En revanche les démarches administratives pour faire admettre au remboursement ces nouvelles technologies par les autorités françaises sont compliquées et elles prennent du temps.

Question d'O Goeau-Brissonnière

Qu'en est-il du non remboursement des techniques de radio fréquence dans les cures de varices ?

Réponse

La radiofréquence dans les cures de varices vient d'être évaluée par l'HAS, ensuite la procédure va se poursuivre auprès de la CNMATS pour la fixation du prix de l'acte.

Question de Fr Gayral

Une étude randomisée menée par mon équipe sur l'effet de la Somatuline dans les fistules digestives a été acceptée dans le monde entier sauf en France ! Qu'en pensez-vous ?

Réponse

Nous ne pouvons regretter que l'essai mené par son équipe n'ait pas pu permettre de faire accepter cette thérapeutique dans le cadre du traitement des fistules digestives.

## Références

1. Activity report of the National committee for the evaluation of medical devices and health technologies. [http://www.has-sante.fr/portail/upload/docs/application/pdf/2012-10/racnedimts\\_2011.pdf](http://www.has-sante.fr/portail/upload/docs/application/pdf/2012-10/racnedimts_2011.pdf)
2. Stordeur S, Vinck I, Neyt M, Van Brabant H, Hulstaert F. Introduction of innovative high-risk medical devices in Europe: are clinical efficacy and safety guaranteed? *Rev Epidemiol Sante Publique*. 2013;61:105-10.
3. Storz-Pfennig P, Schmedders M, Dettloff M. Trials are needed before new devices are used in routine practice in Europe. As the

EU debates new legislation to regulate medical devices. *BMJ*. 2013;346:f1646.

4. Bernard A. The clinical development of new medical devices. *E-Mem Acad Natle Chir* 2012;11:027-31.
5. Boutron I, Moher D, Tugwell P, Giraudeau B, Poiraudeau S, Nizard S et al. A checklist to evaluate a report of nonpharmacological trial (CLEAR NPT) was developed using consensus. *J Clin Epidemiol* 2005;58:1233-40.
6. Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC et al. Challenges in evaluating surgical innovation. *Lancet* 2009;374:1097-104.
7. Vinck I, Neyt M, Thiry N, Louagie M, Ghinet D et al. Introduction of emerging medical devices on the market: a new procedure in Belgium. *Int J Technol Assess Health Care*. 2007;23:449-54.
8. Boutron I, Tubach F, Giraudeau B, Ravaud P. Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. *J Clin Epidemiol* 2004;57:543-50.
9. Kinnart P. Placebo and placebo effect: ethical aspects. *Rev Med Brux* 2007;28:39-44.
10. Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med* 2007;4:61.
11. S Rudicel S, J Esdaile J. The randomized clinical trial in orthopaedics: obligation or option? *J Bone Joint Surg Am* 1985;67:1284-93.
12. Zelen M. A new design for randomized clinical trials. *N Engl J Med* 1979;300:1242-5.
13. Zelen M. Innovations in the design of clinical trials in breast cancer. *Breast Cancer Res Treat* 1983;3:137-42.
14. Zelen M. Strategy and alternate randomized designs in cancer clinical trials. *Cancer treat Rep* 1982;66:1095-100.
15. Zelen M. Randomized consent designs for clinical trials: an update. *Stat Med* 1990;9:645-56.
16. RI Horwitz RI, AR Feinstein AR. Advantages and drawbacks of the Zelen design for randomized clinical trials. *J Clin Pharmacol* 1980;20:425-7.
17. Feng K, Yan J, Li X, Xia F, Ma K, Wang S, Bie P, Dong J. A randomized controlled trial of radiofrequency ablation and surgical resection in the treatment of small hepatocellular carcinoma. *J Hepatology* 2012;57:794-802.
18. Bednarska E, Bryant D, Devereaux PJ. Orthopaedic surgeons prefer to participate in expertise-based randomized trials. *Clin Orthop Relat Res* 2008;466:1734-44.
19. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S et al. Need for expertise based randomized controlled trials. *BMJ* 2005;330:88-91.
20. Mastracci TM, Clase CM, Devereaux PJ, Cina CS. Open versus endovascular repair of abdominal aortic aneurysm: a survey of Canadian vascular surgeons. *J Can Surg* 2008;51:142-8.
21. MacDermid JC, Holtby R, Razmjou H et al. All-arthroscopic versus mini-open repair of small or moderate-sized rotator cuff tears: a protocol for a randomized trial. *BMC Musculoskelet disord* 2006;7:25.
22. Lilford RJ, Braunholz DA, Greenhalgh R, Edwards SJL. Trials and fast changing technologies: the case for tracker studies. *BMJ* 2000;320:43-6.
23. Giraudeau B, Ravaud P. Preventing bias in cluster randomized trials. *PLoS Med* 2009;6:e1000065. doi: 10.1371/journal.pmed.1000065. Epub 2009 May 5.
24. Campbell MJ. Cluster randomized trials in general practice research. *Stat Methods Med Res* 2000;9:81-94.
25. Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2002;1:140-9.
26. Hills M, Armitage P. The two-period cross-over clinical trial. *Br J Clin Pharmacol*. 1979;8:7-20.
27. Armitage P. *Sequential medical trials*. 2nd ed New York Wiley. 1975.
28. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977;64:191-9.
29. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;35:549-56.
30. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. *Br J Cancer* 1976;34:585-612.
31. Falissard B, Lellouch J. A new procedure for group sequential analysis in clinical trials. *Biometrics* 1992;48:373-88.
32. Clayton DG. Ethically optimised designs. *Br J Clin Pharmacol* 1982;13:469-80.
33. Van der Lee JH, Wesseling J, Tanck MWT, Offringa M. Efficient ways exist to obtain the optimal sample size in clinical trials in rare disease. *J Clin Epidemiol* 2008;61:324-30.

34. Sebillé V, Bellissant E. Comparison of four sequential methods allowing for early stopping of comparative clinical trials. *Clin Sci* 2000;98:569-78.
35. Bauer P, Brannath W. The advantages and disadvantages of adaptive designs for clinical trials. *Drug Discov Today* 2004;9:351-7.
36. Simon R. Adaptive treatment assignment methods and clinical trials. *Biometrics* 1977;33:743-9.
37. Lellouch J, Schwartz D. Therapeutic trial: individual or collective ethics. *Rev Inst Int Stat* 1971;39:127-36.
38. Guidance for the use of bayesian statistics in medical device clinical trial. Center for devices and radiological health, center for biologics evaluation and research. Silver Spring, Rockville: CDRH, CBER 2006. <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationGuidance/GuidanceDocumentation/ucmo071121.pdf>
39. Berry SM, Carlin BP, Lee JJ, Muller P. Bayesian adaptive methods for clinical trials. Ed Chapman & Hall/CRC Biostatistics series 2011.
40. Campbell G. Bayesian statistics in medical devices: innovation sparked by the FDA. *J Biopharm Stat* 2011;21:871-87.
41. Ribouleau L, Chevret S. Bayesian statistical method was under-used despite its advantages in the assessment of implantable medical devices. *J Clin Epidemiol* 2011;64:270-9.
42. Wilber DJ, Pappone C, Neuzil P, De Paola A, Marchlinski F, Natale A et al Comparison of antiarrhythmic drug therapy and radiofrequency catheter ablation in patients with paroxysmal atrial fibrillation. A randomized controlled trial. *JAMA* 2010;303:333-40.
43. Kearney RS, Achten J, Parsons NR, Costa ML. The comprehensive cohort model in a pilot trial in orthopaedic trauma. ML Costa. *BMC Med Res Methodol* 2011;11:39.
44. Torgerson DJ, Sibbald B. Understanding controlled trials. What is a patient preference trial? *BMJ* 1998;316:360.
45. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* 1984;79:516-24.
46. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41-55.
47. Yue LQ. Statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. *J Biopharm Stat* 2007;17:1-13.
48. Stuart EA. Matching methods for causal inference: a review and a look. *Stat Sci* 2010;25:1-21.
49. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399-424.
50. Rosenbaum PR. Discussing hidden bias in observational studies. *Ann Intern Med* 1991;115:901-5.
51. Dahabreh IJ, Sheldrick RC, Paulus JK, Chung M, Varvarigou V, Jafri H, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J* 2012;33:1893-901.