



HAUTE AUTORITÉ DE SANTÉ

**GUIDE MÉTHODOLOGIQUE**

**Méthodes quantitatives pour évaluer les  
interventions visant à améliorer les pratiques**

**Juin 2007**

Ce guide méthodologique est téléchargeable sur  
[www.has-sante.fr](http://www.has-sante.fr)

Haute Autorité de santé  
Service communication  
2 avenue du Stade de France – F 93218 Saint-Denis La Plaine  
CEDEX  
Tél. : +33 (0)1 55 93 70 00 – Fax : +33 (0)1 55 93 74 00

Ce document a été validé par le Collège de la Haute Autorité de santé en juin 2007.

© Haute Autorité de santé – 2007

---

# Sommaire

<b>Introduction .....</b>	<b>3</b>
<b>Méthodes quantitatives pour évaluation de l'impact d'interventions visant à améliorer les pratiques.....</b>	<b>5</b>
<b>1. Les essais contrôlés randomisés .....</b>	<b>7</b>
1.1 Essais randomisés à unité de randomisation individuelle.....	7
1.2 Essais randomisés à unité de randomisation collective (essais en cluster, essais en grappes).....	7
<b>2. Études quasi expérimentales .....</b>	<b>16</b>
2.1 Séries chronologiques.....	16
2.2 Études avant/après avec site contrôle contemporain de l'intervention .....	20
<b>3. Autres particularités méthodologiques des études d'intervention .....</b>	<b>22</b>
3.1 Standardisation de l'intervention .....	22
3.2 Critères de jugement ou indicateurs.....	22
3.3 Aveugle.....	23
3.4 Sources de données .....	23
<b>Concevoir et mettre en œuvre une étude.....</b>	<b>25</b>
<b>1. Avant la réalisation de l'étude .....</b>	<b>25</b>
1.1 La formulation du problème.....	26
1.2 Détermination de la population-cible .....	29
1.3 Méthodes de recrutement .....	30
1.4 Choix de la stratégie d'amélioration .....	30
1.5 Choix du plan d'étude expérimental ou d'observation .....	32
1.6 Choix des indicateurs ou critères de jugement .....	33
1.7 Aspects éthiques .....	36
1.8 Enregistrement de l'étude .....	36
<b>2. Durant la réalisation de l'étude.....</b>	<b>37</b>
2.1 Maintien de l'intérêt des participants après le recrutement.....	37
2.2 Sécurité et effets indésirables .....	37
2.3 Gestion des données .....	38

<b>3. Après la réalisation de l'étude.....</b>	<b>39</b>
3.1 Rapport final.....	39
3.2 Politique de publication .....	40
<b>Références .....</b>	<b>41</b>
<b>ANNEXE 1 : <i>Cochrane Effective Practice and Organisation of Care Group</i></b> <b>(EPOC).....</b>	<b>43</b>
<b>ANNEXE 2 : Grille de choix des thèmes de travail .....</b>	<b>46</b>
<b>ANNEXE 3 : Analyse critique d'une étude d'intervention .....</b>	<b>47</b>
<b>Fiche descriptive .....</b>	<b>55</b>

# Introduction

La Haute Autorité de Santé remercie le Pr Philippe RAVAUD et le Dr Pierre DURIEUX, auteurs de ce guide.

Les interventions destinées à améliorer la qualité des soins sont trop souvent mises en place sans prévoir l'évaluation de leur impact ou en évaluant cet impact, avec une méthodologie dont on sait d'emblée qu'elle ne permettra pas de répondre correctement à la question. Sur les quelques centaines d'études d'interventions référencées par le groupe de travail EPOC (*Effective Practice and Organization of Care*) de la collaboration *Cochrane*, moins d'une dizaine ont été réalisées en France (1,2). Ceci est également vrai pour les politiques visant à modifier les comportements, les mises en place, sans que l'évaluation de leur impact ait été sérieusement envisagée (3,4).

Évaluer correctement les interventions destinées à améliorer la qualité des soins et à modifier les comportements est pourtant essentiel, afin de savoir quelles interventions sont susceptibles d'être efficaces dans notre système de santé, en contribuant ainsi d'une part à l'amélioration de la qualité des soins, et d'autre part à l'amélioration de notre politique de santé. Ces interventions sont fondées le plus souvent sur une modification de la pratique professionnelle ou de l'organisation des soins. Lorsqu'elles sont en phase de développement, elles doivent faire, à l'instar de toute innovation, l'objet d'une évaluation de ses effets positifs et négatifs, avant d'être généralisées. Cette évaluation est fondée principalement sur des méthodes quantitatives, domaine de l'épidémiologie et des biostatistiques. Cependant, les méthodes qualitatives, issues des sciences humaines et sociales, sont actuellement en plein développement, pour aider à la prise en compte des freins à l'amélioration des pratiques, pour faciliter le choix d'une stratégie d'intervention adaptée ou pour analyser les raisons de l'échec d'une intervention. Ce document présente les approches quantitatives uniquement.

Les méthodes proposées dans ce document sont appliquées dans un cadre de recherche, et ne sont pas nécessaires lors de l'utilisation routinière de ces interventions. Leur utilisation contribue, cependant, à faire en sorte que les interventions mises en place pour améliorer les pratiques médicales aient fait l'objet d'une évaluation avant leur utilisation routinière.

L'objectif de ce document est, à partir des données de la littérature médicale scientifique et de l'expérience de ses auteurs, de présenter les différentes méthodes quantitatives d'évaluation des interventions, destinées à améliorer la qualité des soins. Il en décrit opérationnellement les différentes étapes de mise en œuvre.

La stratégie de recherche documentaire a donc été ciblée sur la méthodologie de ces études. Les bases bibliographiques des auteurs/experts du rapport ont été exploitées. Ces bases sont alimentées par une veille documentaire systématique :

- dans *PubMed/Medline*, à partir des termes :  
*Cluster randomised controlled OU trial time series analysis OU time-series intervention analysis OU analysis of interrupted time series OU Interrupted time series design OU interrupted time-series analysis OU interrupted time series studies OU alternating time series analysis OU controlled "before-after" study OU pre/post study ;*
- ainsi que par le dépouillement des revues suivantes :  
*BMJ, The Lancet, Annals of Internal Medicine, JAMA, Statistics in Medicine, BMC Health Services Research, American Journal of Public Health, International Journal of Technology Assessment in Health Care, Quality Safety Health Care, International journal for quality in health care.*

Ce guide constitue avant tout un **cadre de référence pour les chercheurs**. Il vise à soutenir le développement actuel de la recherche sur les systèmes de santé, soutenue notamment par les **appels à projets de recherche de la HAS, de la CNAMTS, de la DHOS** (appel à projet PrEQHos). Il est une aide aux décideurs et professionnels de santé appelés à expertiser les projets de recherche portant sur une évaluation d'intervention.

Ce guide constitue également un **document de référence pour la lecture critique** utile aux professionnels de terrain. Il fait cependant référence à des concepts, méthodes et outils complexes qui ne sont pas développés, et qui peuvent nécessiter l'aide de méthodologistes. Nous espérons cependant qu'à sa lecture, le professionnel de santé soit sensibilisé à l'importance de l'évaluation rigoureuse d'une intervention innovante, qui conditionnera peut-être sa pratique de demain.

## Méthodes quantitatives pour évaluation de l'impact d'interventions visant à améliorer les pratiques

Lorsque l'on met en place une intervention destinée à modifier une pratique, et améliorer la qualité des soins, il apparaît logique de mesurer l'état de cette pratique avant et après la mise en place de l'intervention, et de conclure à l'efficacité de l'intervention si la pratique s'est améliorée entre les deux mesures.

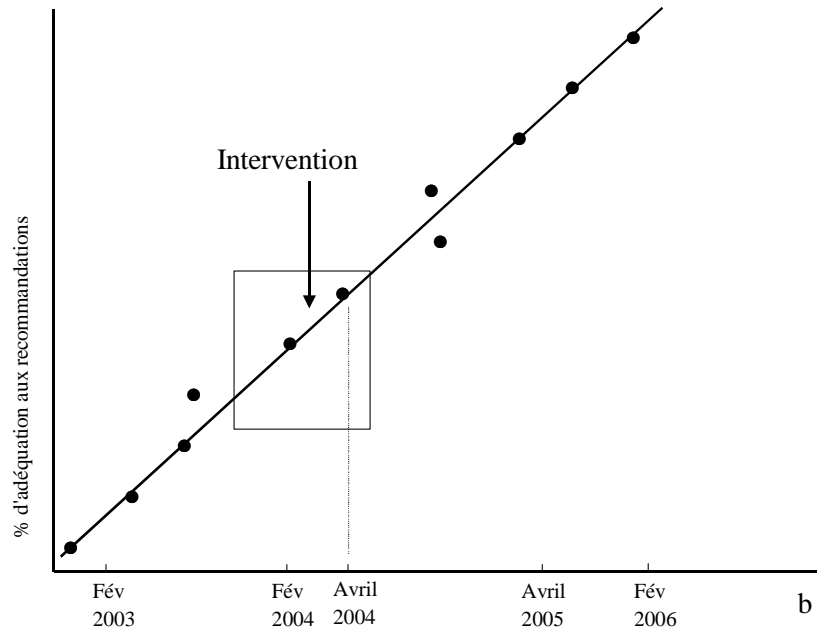
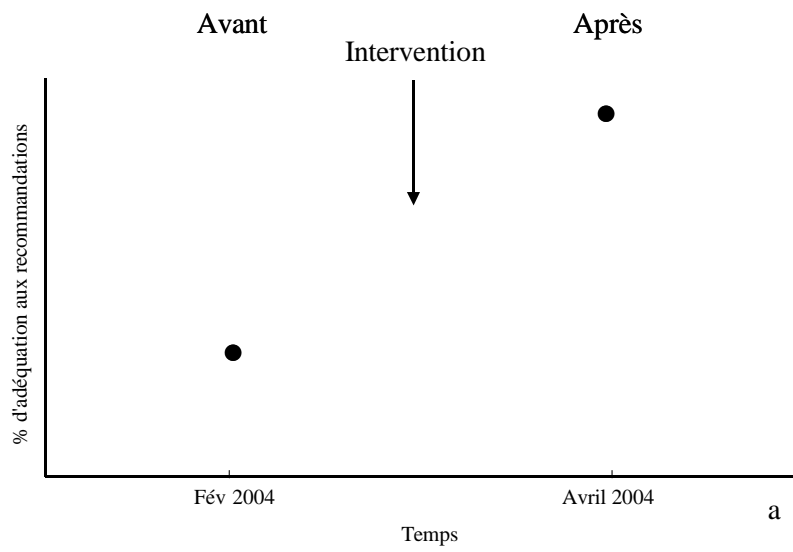
En réalité, en raison de tendances préexistantes à l'amélioration ou de variations liées à une autre cause que l'intervention, il est difficile, voire impossible, d'attribuer les variations observées à l'intervention (*figure 1*). Il a de plus été suggéré que les résultats des études avant/après pouvaient surestimer les effets des interventions ayant pour but d'améliorer la qualité des soins (5).

Les études avant/après, très (trop) souvent réalisées, ne devraient donc pas être utilisées pour évaluer l'impact d'une intervention. Ce type d'étude permet simplement de montrer l'évolution d'une pratique, sans que l'on puisse affirmer que les modifications observées sont liées à l'intervention.

Évaluant l'impact d'une conférence de consensus sur le cancer de la prostate, Kosecoff *et al.*, en 1987, avaient ainsi montré que les pratiques recommandées (chirurgie dans le cancer localisé) s'étaient améliorées après la publication des recommandations de la conférence, mais qu'il s'agissait en réalité d'une tendance à l'amélioration préexistante. Lorsque l'on analysait les données sur une période prolongée, et non pas simplement par un point de mesure avant et un point de mesure après intervention, il était observé que la publication des recommandations n'avait eu aucun impact sur les pratiques (cf. *figure 1*) (6).

Personne ne discute aujourd'hui que l'essai contrôlé randomisé constitue la référence pour évaluer l'efficacité d'un nouveau médicament. Il en est aujourd'hui de même pour toute intervention sur la santé. Dans le cas de l'évaluation d'un médicament, la mesure des résultats de santé se fait sur la cible de l'intervention (dans les deux cas, il s'agit du patient). L'une des spécificités de l'évaluation des interventions d'amélioration des pratiques et de changement des comportements est que la cible de l'intervention est collective (un groupe de médecins, d'établissements, etc.), alors que la mesure des résultats de santé est à l'échelle des patients. Cette spécificité fait que l'essai contrôlé randomisé sera le plus souvent réalisé en clusters (ou en grappes). C'est pourquoi nous développons longuement ce schéma expérimental de référence. Cependant, des schémas dits quasi expérimentaux (étude avant/après avec groupe-contrôle, série chronologique) peuvent être envisagés lorsque des barrières politiques, éthiques ou organisationnelles empêchent la réalisation d'un essai.

**Figure 1.** Étude avant/après. Les variations observées entre février 2004 et avril 2004 (partie a) peuvent n'être liées qu'à la tendance préexistante.





# 1. Les essais contrôlés randomisés

## 1.1 Essais randomisés à unité de randomisation individuelle

L'essai contrôlé randomisé représente la méthode de référence pour évaluer l'impact d'une intervention en santé. Il permet d'évaluer l'impact de cette intervention, à l'aide d'une comparaison directe avec un groupe qui ne reçoit pas l'intervention ou bien qui reçoit une autre intervention.

La forme la plus classique de ces essais est l'essai à unité de randomisation individuelle, c'est-à-dire un essai où les malades sont randomisés, et se voient attribuer l'intervention A ou l'intervention B. La randomisation permet d'obtenir que tous les facteurs pronostiques pouvant influencer sur l'effet de l'intervention, qu'il s'agisse de facteurs pronostiques connus mais également inconnus, soient équitablement répartis entre les deux groupes de l'essai. Ce point est d'autant plus important que l'on connaît encore moins bien pour les essais thérapeutiques les facteurs de confusion potentiels, liés à l'organisation des soins ou aux performances professionnelles. Il est donc très difficile d'envisager un ajustement sur ces paramètres.

Les essais à unité de randomisation individuelle sont rarement utilisés dans le domaine de l'évaluation des interventions visant à améliorer la qualité des soins. Le problème majeur de ces essais est en effet le risque de « contamination » entre les groupes. Ainsi, si l'on désire comparer deux interventions, destinées à des médecins et ayant pour but d'améliorer la qualité des soins (par exemple, mise en œuvre de recommandations *versus* absence de mise en œuvre de ces recommandations), il est impossible d'imaginer qu'un médecin va appliquer une recommandation pour le malade affecté de manière aléatoire à l'intervention A, puis oublier l'existence de cette recommandation pour le malade affecté de manière aléatoire à l'intervention B. Dans le domaine de la qualité de soins, les essais randomisés à unité de randomisation individuelle sont donc rarement utilisés, sauf dans des cas très particuliers. Ce type d'essai a par exemple été utilisé pour évaluer la mise en œuvre de recommandations de pratiques dans un système de prescriptions médicamenteuses informatisées. Le médecin recevant un rappel au moment de sa prescription pour certains patients et pas pour d'autres, le risque de contamination apparaissait dès lors faible (7).

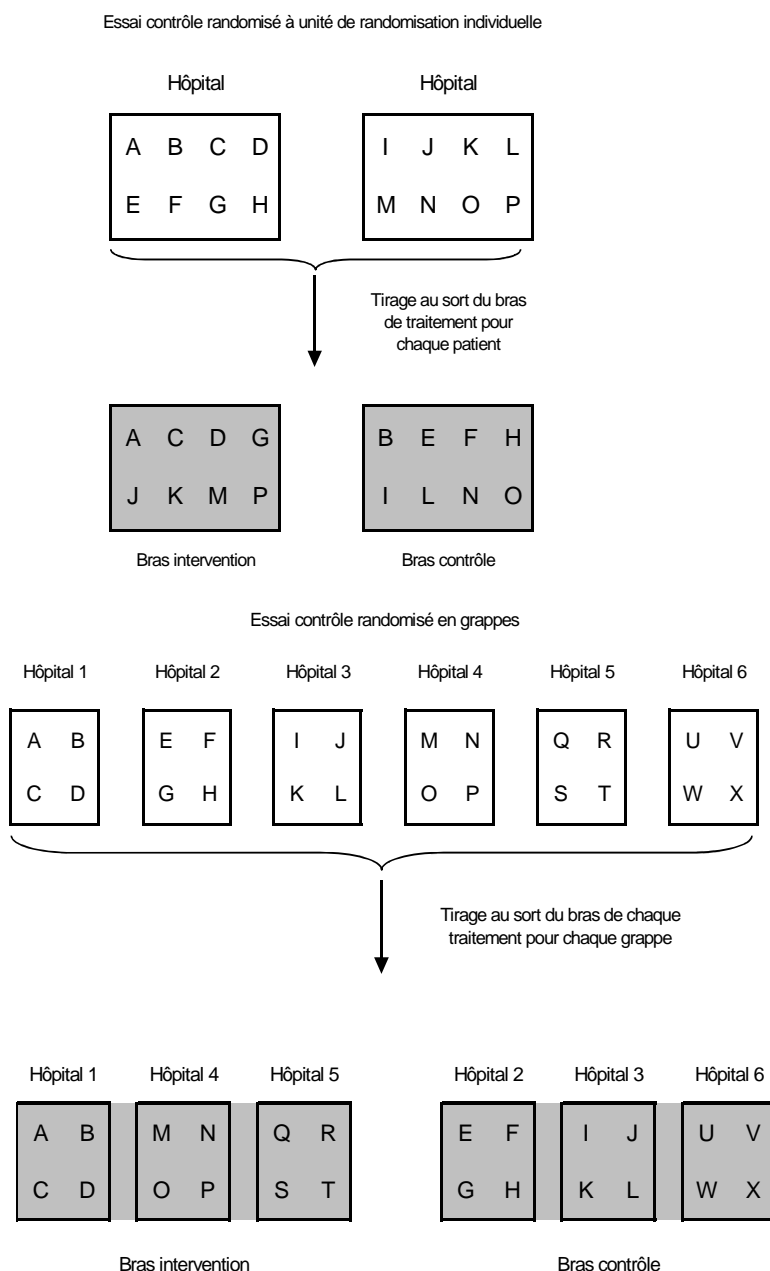
L'hypothèse de base dans un essai randomisé à unité de randomisation individuelle est que les différents malades sont indépendants. Cette hypothèse est violée dans la mise en œuvre d'interventions destinées à améliorer la qualité des soins. En effet, les malades suivis par un même médecin ou dans un même hôpital (à qui l'intervention est destinée) ont une probabilité plus élevée d'avoir la même prise en charge ou de répondre de la même manière à l'intervention, que des malades traités par des médecins différents ou dans des hôpitaux différents. Il en résulte une perte de puissance, dont on doit tenir compte à la fois dans le schéma de l'étude, dans le calcul du nombre de sujets nécessaires et dans l'analyse statistique. La réponse à ces problèmes est la réalisation d'un essai randomisé à unité de randomisation collective, appelé essai en cluster ou en grappes (l'ensemble des malades suivis par un même médecin ou par une même structure de soins constitue une grappe).

## 1.2 Essais randomisés à unité de randomisation collective (essais en cluster, essais en grappes)

Dans ce type d'essai, ce ne sont donc plus des individus qui sont randomisés, mais des groupes d'individus (hôpitaux, services hospitaliers, unités de soins, médecins, infirmières, etc. [cf. *figure 2*]), ce qui est finalement logique, puisque c'est à eux que l'intervention est destinée. La randomisation a donc lieu à un niveau collectif (médecins,

services, etc.), et les données sont recueillies généralement à un niveau différent (malade).

**Figure 2.** Comparaison entre le mode de recrutement des individus et de randomisation, dans les essais à unité de randomisation individuelle et dans ceux à unité de randomisation collective



L'avantage majeur des essais en cluster est de limiter ou d'éviter le risque de contamination, c'est-à-dire le risque, par exemple que les médecins randomisés dans le bras-témoin soient influencés par les professionnels du bras-intervention, et améliorent leurs pratiques. La contamination peut entraîner une sous-estimation de l'effet réel de l'intervention, et augmenter le risque d'erreur de type II. Le nombre de cluster étant souvent faible, il est souvent utile de prendre en compte la performance initiale, avant intervention.

### 1.2.1 Particularités statistiques des essais à unité de randomisation collective

La mesure de l'étendue de l'effet cluster ou grappes est appelée « Coefficient de corrélation intraclasse » (CCI). Ce coefficient évalue le degré de ressemblance des individus au sein d'une grappe, et est basé sur la relation entre la variance entre clusters et la variance intracluster. Pour cette raison, les tailles des échantillons doivent être « majorées » pour tenir compte de l'effet de mise en grappes. L'augmentation requise est influencée à la fois par le CCI et par la taille des grappes. On doit augmenter la taille standard de l'échantillon par un facteur d'inflation égal à  $[1 + (n - 1) \rho]$ , où  $n$  est la taille moyenne d'une grappe et  $\rho$  le CCI estimé. On peut obtenir le nombre de participants supplémentaires en augmentant soit le nombre de grappes dans l'étude, soit le nombre de participants par grappe. Toutefois, l'augmentation de la taille de la grappe à plus de 50 ne permettra d'améliorer que légèrement l'efficacité statistique, l'augmentation du nombre de grappes sera beaucoup plus efficace (8). Il faut ainsi faire un choix entre les difficultés logistiques et les coûts associés au recrutement de nouvelles grappes, et les difficultés associées à l'augmentation du nombre de patients par grappe.

Si le nombre de grappes est faible ( $< 5$ ), quelque soit la taille de chaque grappe, on peut dans certaines circonstances ne pas atteindre une puissance de 80 % (cf. *figure 3*).

Le coefficient de corrélation intraclasse varie selon le type de critères de jugement utilisé. Plus le CCI est élevé, plus l'effet cluster sera important (un CCI à 0 signifie qu'il n'existe aucun effet cluster). Les valeurs du CCI sont plus élevées pour les critères de jugement de type procédure (valeur médiane 0,063, écart interquartile : 0,016 à 0,124) que pour les critères de jugement de type résultat (valeur médiane 0,030, écart interquartile 0,005 à 0,052). Ces CCI varient probablement aussi selon d'autres paramètres : le contexte de soins (CCI plus faible en soin primaire), l'unité de randomisation, le type de mesure subjective ou objective (9-12). Pour les calculs d'effectifs, il est nécessaire de faire des hypothèses à propos du CCI retenu. Ces hypothèses seront faites en fonction des paramètres connus pour influencer sur le CCI, des CCI retrouvés dans des études précédemment publiées, des résultats d'études préliminaires éventuelles. Il est prudent de faire varier le CCI, car des erreurs sur ces estimations initiales sont fréquentes. À la fin de l'étude, il est indispensable, surtout en cas d'étude négative, de calculer le CCI réel afin de vérifier que l'étude avait une puissance suffisante.

Un exemple de calcul de nombre de sujets est donné dans l'encadré n°1.

Une autre difficulté, avec les essais à unité de randomisation collective, est le risque de déséquilibre ou de non-comparabilité entre les groupes. Si dans les essais à unité de randomisation individuelle ce risque est faible, car le nombre d'unités de randomisation est élevé (fréquemment supérieur à 100 ou 200), ce risque est majeur dans les essais à unité de randomisation collective, où l'on randomise parfois moins de 10 unités (hôpitaux, structures de soins ou médecins).

Bien évidemment, plus le nombre d'unités de randomisation disponible est faible, plus le risque d'un déséquilibre dû à la chance entre les groupes-témoins et les groupes-intervention est élevé.

Une difficulté supplémentaire dans ces essais est représentée par le biais induit à différent niveau par l'absence d'aveugle. Il n'est pas rare dans les études en cluster que les hôpitaux, services ou médecins qui se voient assigner une intervention qui ne leur convient pas (par exemple s'ils sont randomisés groupe-témoins) refusent finalement de participer à l'étude induisant bien évidemment un biais, et aggravant le risque de déséquilibre entre les groupes.

Enfin l'analyse des essais en cluster est plus complexe que l'analyse des essais à unité de randomisation individuelle. Fréquemment, ces essais sont analysés à l'échelon individuel sans prendre en compte l'effet du cluster, ce qui est incorrect au plan statistique. Ces essais peuvent être analysés à l'échelon du cluster en utilisant le cluster

comme unité de randomisation mais aussi d'analyse. Dans ce cas, une statistique résumée (moyenne ou pourcentage) est obtenue pour chaque cluster.

Ces données disponibles pour chaque cluster peuvent être considérées comme indépendantes, et analysées en utilisant les tests statistiques usuels. La deuxième méthode qui peut être utilisée est l'analyse des données à l'échelon individuel, en réalisant un ajustement pour prendre en compte l'effet cluster. La dernière méthode consiste à utiliser des techniques de modélisation prenant en compte la nature hiérarchique des données. Ainsi, pour un essai où sont randomisés des services hospitaliers, on peut utiliser un modèle où les malades sont « emboîtés » dans le service hospitalier, et prenant en compte des covariables mesurées à l'échelon du malade (âge, sexe, gravité de la maladie) ou du service hospitalier (taille, nombre de médecins, etc.).

Par exemple, afin d'évaluer l'impact de la mise en œuvre d'une intervention destinée à utiliser des règles de prescription des radiographies de la cheville, chez des patients se présentant aux urgences de l'Assistance publique – Hôpitaux de Paris pour traumatisme de la cheville (règles d'Ottawa), deux hôpitaux ont été randomisés « intervention », et trois hôpitaux ont été randomisés « pratique usuelle ». Le pourcentage de réalisation de radiographies de cheville a été évalué avant et après intervention dans les deux groupes d'hôpitaux.

Dans cette étude, l'unité de randomisation était bien l'hôpital et les données étaient recueillies auprès des patients, mais il fallait aussi prendre en compte un autre niveau de prise en charge, le médecin (1 005 patients pris en charge par 44 médecins dans le groupe-intervention, 906 patients pris en charge par 47 médecins dans le groupe-témoins). L'analyse statistique, en comparant les pourcentages de réalisation des radiographies chez les patients issus des deux groupes d'hôpitaux, a ainsi pris en compte à la fois l'effet hôpital et l'effet médecin (en donnant plus de poids aux médecins qui avaient vu plus de malades), c'est-à-dire deux niveaux de grappes (13).

## 1.2.2 Types d'essais en cluster

### 1.2.2.1 Essais comparant deux types d'interventions

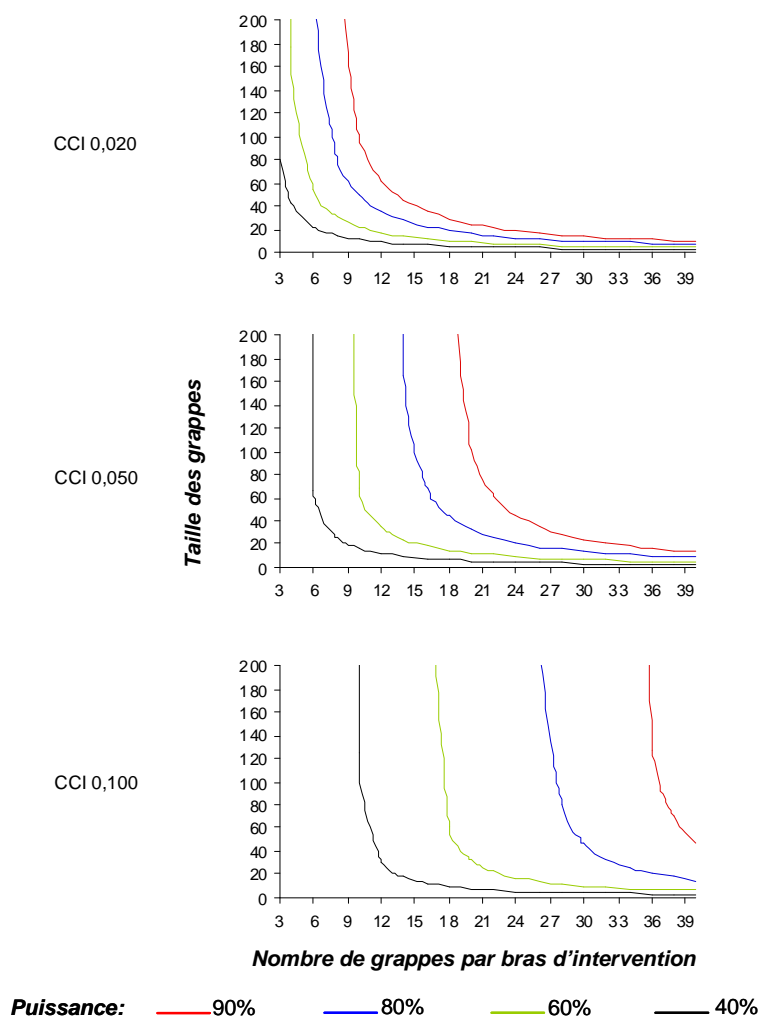
Le schéma le plus simple est l'essai comparant deux types d'intervention, comme dans l'exemple précédent où l'on randomise un groupe-intervention et un groupe-témoins. Ces essais sont relativement simples en termes de mise en œuvre et d'analyse. Cependant, ils fournissent uniquement des informations sur l'efficacité d'une seule intervention en la comparant à un groupe-témoin (ou sur l'efficacité relative de deux interventions sans comparaison à un groupe-témoins).

Dans le cas d'une comparaison entre une intervention et l'absence d'intervention, ce type d'essai expose au risque dit d'effet Hawthorne : l'effet positif observé dans le groupe-intervention peut être lié au fait que les sujets randomisés dans ce groupe améliorent leur pratique simplement parce qu'ils se sentent observés (ou que les sujets-témoins sont démotivés par le simple fait d'être dans le groupe-témoins).

### 1.2.2.2 Essai comparant des interventions multiples

Ce type d'essai est une simple extension de l'essai comparant deux interventions (par exemple un groupe-intervention A, un groupe-intervention B, un groupe-témoins). Ce type d'étude a, par exemple, été proposé par des auteurs cherchant à comparer deux niveaux d'intensité d'une même intervention à l'absence d'intervention. Cependant, il faut prendre en compte la perte de puissance statistique. Par exemple, pour un essai comparant 3 interventions, il est nécessaire d'augmenter la taille de l'échantillon de 50 % pour atteindre la même puissance que dans un essai comparant deux interventions.

**Figure 3.** Influence de la taille et du nombre de grappes sur la puissance pouvant être obtenue.



## **Encadré n°1 – Exemple de calcul d'effectif pour les essais à unité de randomisation individuelle ou collective**

Horbar JD, Carpenter JH, Buzas J, Soll RF, Suresh G, Bracken MB, et al. Collaborative quality improvement to promote evidence based surfactant for preterm infants : a cluster randomised trial. *BMJ* 2004;329(7473):1004.(14).

*Objectif* : évaluer une intervention multifacette d'amélioration de la qualité, pour promouvoir l'utilisation d'un traitement par surfactant chez les enfants prématurés âgés de 23 à 29 semaines.

L'effectif nécessaire sera extrêmement différent selon le plan expérimental retenu (essai à randomisation individuelle ou collective), le type de critère de jugement :

1- Si, essai randomisé à unité de randomisation individuelle (randomisation de malades)

-  $\alpha = 0,05$  ;  $\beta = 0,20$ .

- Critère de jugement de type procédure : % d'enfant recevant du surfactant donné en salle de délivrance.

- Effet attendu de l'intervention (20 % dans le bras cohorte, 50 % dans le bras intervention).

Le nombre nécessaire de malades est de 78.

2- Si, essai à unité de randomisation collective

L'effectif nécessaire augmentera, car il faudra prendre en compte un facteur d'inflation,  $FI = 1 + (m-1)\rho$ , où  $m$  est la taille des grappes (nombre d'individus vu par un médecin par exemple si l'on randomise des médecins), et  $\rho$  est le coefficient de corrélation intraclasse qui quantifie la corrélation entre les malades d'une même grappe. Si le facteur d'inflation est égal à 10, il faudra 10 fois plus de malades dans un essai à unité de randomisation collective que dans un essai à randomisation individuelle pour obtenir une même puissance. La valeur du coefficient de corrélation intraclasse dépend en particulier de l'unité de randomisation (les malades suivis par un même médecin sont probablement plus ressemblants que les malades suivis au sein d'un hôpital) et du type de critère de jugement. Les critères de jugement de type pratique médicale sont probablement plus corrélés au sein d'une grappe (par exemple parmi les malades traités par un même médecin) que les critères de jugement de type résultats, par exemple la mortalité.

a – unité de randomisation : service.

- Critère de jugement de type procédure (surfactant donné en salle de délivrance).

- Coefficient de corrélation intraclasse estimé (CCI) 0,31.

Le CCI reflète le degré de corrélation ou de ressemblance entre les prescriptions d'un même service (ou d'un même médecin).

Si on a 20 services disponibles au total, on ne pourra jamais atteindre une puissance de 80 %, quelque soit le nombre d'enfants inclus dans chaque service (puissance maximum 72 %).

Si on a 24 clusters, chacun incluant 375 sujets (9 000 enfants au total), on obtient une puissance de 80 %.

Si on a 30 services, il suffit que chacun inclut 9 sujets (270 enfants au total) pour obtenir une puissance de 80 %.

Il apparaît donc clairement que la puissance de l'essai est plus liée au nombre de services randomisés qu'au nombre de malades inclus dans chaque service.

b – unité de randomisation : médecins et critère de jugement de type procédure.

- CCI = 0,50 (logiquement plus élevé que dans le cas d'une randomisation par service).

Si on a moins de 38 clusters (médecins), la puissance maximale atteinte sera inférieure à 80 % quelque soit le nombre de malades vus par chaque médecin.

Si on a 50 médecins randomisés incluant chacun 4 malades (soit 200 malades au total), la puissance sera de 81,6 %.

Si on a 60 médecins randomisés incluant chacun 2 malades (soit 120 malades au total), la puissance sera de 80 %.

c – unité de randomisation : médecins et critère de jugement de type résultats (mortalité).

- CCI 0,01 (le CCI est toujours plus bas pour un critère de type résultat que pour un critère de type procédure).

Il est nécessaire pour mettre en évidence une réduction de la mortalité de 30 % (en prenant l'hypothèse que la mortalité est de 20 % dans le groupe-témoins), avec une puissance de 80 %, d'avoir 20 clusters (médecins) de 42 malades (840 malades au total) ou 40 clusters de 17 malades (680 malades au total).

### 1.2.2.3 Plan factoriel

Les essais cliniques factoriels évaluent simultanément l'effet de deux interventions ou plus, en utilisant diverses combinaisons de ces interventions. Le cas le plus simple est le plan factoriel 2 x 2, dans lequel les grappes sont réparties aléatoirement dans l'une des quatre combinaisons de deux interventions (disons A et B). Ces combinaisons sont A seul, B seul, A et B ; ni A ni B (témoin).

Ce plan permet de comparer chacune des interventions expérimentales avec le témoin, de comparer les interventions expérimentales entre elles, et d'étudier les interactions possibles entre les deux interventions (c'est-à-dire, de comparer la somme des effets de A et B donnés séparément avec les effets de la combinaison).

S'il n'y a pas d'interaction entre les interventions, le plan factoriel aura une puissance statistique plus importante qu'un essai traditionnel comparant plusieurs types d'intervention. Ce plan permet d'étudier deux interventions dans le même essai, sans augmenter le nombre de grappes. L'inconvénient est qu'il est souvent très difficile de garantir qu'il n'y aura pas d'interaction, et par conséquent l'interprétation des résultats peut être difficile. Ces essais sont également plus difficiles à concevoir, à mettre en œuvre et à analyser.

Un exemple de ce type est donné dans l'encadré n°2.

#### **Encadré n°2 – Exemple de plan factoriel**

*Thomas RE, Croal BL, Ramsay C, Eccles M, Grimshaw J. Effect of enhanced feedback and brief educational reminder messages on laboratory test requesting in primary care: a cluster randomised trial. Lancet 2006;367(9527):1990-6.(15).*

**Objectif :** évaluer l'effet de deux interventions sur le nombre d'examens de laboratoire prescrits par des médecins généralistes.

**Plan expérimental :** essai contrôlé randomisé en cluster, avec plan factoriel 2 x 2.

Groupe 1 Aucune intervention (témoin)	Groupe 2 Reçoit retour d'informations
Groupe 3 Reçoit messages éducationnels	Groupe 4 Reçoit retour d'informations + messages éducationnels

**Participants :** 85 cabinets de médecins généralistes du Nord-Est de l'Écosse.

**Interventions :** document diffusé à chaque cabinet de médecins, présentant leurs données de prescription des 3 dernières années concernant 9 tests biologiques comparées aux données régionales (*feedback*); recommandations simples accompagnant chaque résultat d'un des 9 tests biologiques prescrit par un médecin des cabinets d'intervention concernés (message éducationnel).

**Critère de jugement :** nombre d'examens prescrits avant et durant les 12 mois de la période intervention.

**Résultat :** baisse du nombre total d'examens d'environ 10 % dans le groupe retour d'informations, 10 % dans le groupe messages éducationnels et 20 % dans le groupe retour d'informations + messages éducationnels (effet additif des deux types de messages).

#### 1.2.2.4 Modèle équilibré incomplet

Le schéma dit en blocs équilibrés incomplets est fréquemment utilisé pour éviter l'effet Hawthorne décrit plus haut. Le cas le plus simple est un modèle de blocs équilibrés incomplets 2 x 2. Dans ce schéma expérimental, un groupe reçoit une intervention pour la prise en charge du problème de santé ou de la maladie A, et sert de contrôle pour la prise en charge de la maladie B. L'autre groupe reçoit l'intervention pour la maladie B, et sert de contrôle pour la maladie A. Étant donné que les sujets des deux groupes sont soumis au même niveau d'intervention, l'effet Hawthorne doit être égal dans les deux groupes.

Un exemple de ce type de schéma est donné dans l'encadré n°3.

#### **Encadré n°3 – Exemple d'essai randomisé en cluster avec plan expérimental de type bloc incomplet.**

*Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care : cluster randomised controlled trial. BMJ 2002;325(7370):941.(16).*

*Objectif* : évaluer l'impact d'un système d'aide à la décision informatisé, pour la mise en œuvre de recommandations de pratique clinique concernant la prise en charge de l'asthme et de l'angor chez les adultes, dans le cadre des soins primaires.

*Plan expérimental* : essai randomisé pragmatique avec unité de randomisation collective (randomisation de cabinets de médecins) de type bloc incomplet équilibré.

	Asthme	Angine
Médecins groupe 1	Intervention	Témoin
Médecins groupe 2	Témoin	Intervention

*Participants* : 60 cabinets de médecine générale du Nord-Est de l'Angleterre. Médecins généralistes et infirmiers, dont les médecins utilisent des systèmes informatisés pour leur prescription.

*Intervention* : Mise en place de recommandations informatisées permettant de guider la consultation, soit sur l'asthme (médecins du groupe 1), soit sur l'angor (médecins du groupe 2).

*Critère de jugement* : Adhésion aux recommandations, fondée sur l'analyse des dossiers, effectuée avant et après mise en œuvre de l'aide.

CCI estimé de 0,05.

Analyse en intention de traitement, c'est-à-dire que les médecins ont ou non utilisé l'aide informatique

*Résultat* : aucun impact de l'aide informatique sur la qualité de la prise en charge. La non-efficacité de l'aide semble principalement due au fait que les médecins ne la consultaient pas.



## **Pour en savoir plus sur les essais à unité de randomisation collective**

### **Exemples pratiques**

Beck CA, Richard H, Tu JV, Pilote L. Administrative Data Feedback for Effective Cardiac Treatment : AFFECT, a cluster randomized trial. JAMA 2005;294(3):309-17.

Moher M, Yudkin P, Wright L, Turner R, Fuller A, Schofield T, *et al.* Cluster randomised controlled trial to compare three methods of promoting secondary prevention of coronary heart disease in primary care. BMJ 2001;322(7298):1338.

Ornstein S, Jenkins RG, Nietert PJ, Feifer C, Roylance LF, Nemeth L, *et al.* A multimethod quality improvement intervention to improve preventive cardiovascular care : a cluster randomized trial. Ann Intern Med 2004;141(7):523-32.

Perria C. Strategies for the introduction and implementation of a guideline for the treatment of type 2 diabetics by general practitioners (GPs) of the Lazio region of Italy (IMPLEMEG study) : protocol for a cluster randomised controlled trial ISRCTN80116232. BMC Health Serv Res 2004;4(1):13.

Yealy DM, Auble TE, Stone RA, Lave JR, Meehan TP, Graff LG, *et al.* Effect of increasing the intensity of implementing pneumonia guidelines. A randomized trial. Ann Intern Med 2005;143:881-94.

### **Quelques références méthodologiques**

Campbell MJ, Donner A, Klar N. Developments in cluster randomized trials and Statistics in Medicine. Stat Med 2007;26(1):2-19.

Donner A, Klar N. Design and analysis of cluster randomization trials. Arnold : London, 2000.

Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. Am J Public Health 2004;94(3):416-22.

Elbourne DR, Campbell MK. Extending the CONSORT statement to cluster randomized trials : for discussion. Stat Med 2001;20(3):489-96.

Murray DM, Varnell SP, Blitstein JL. Design and analysis of group-randomized trials : a review of recent methodological developments. Am J Public Health 2004;94(3):423-32.

## 2. Études quasi expérimentales

### 2.1 Séries chronologiques

Les études de type séries chronologiques permettent de détecter si une intervention a permis de modifier une tendance préexistante. Ce type d'étude est surtout utile dans deux situations :

- lorsqu'il est impossible de randomiser, car l'intervention ne peut être limitée à une partie des centres, c'est le cas par exemple d'une campagne télévisée visant à réduire le tabagisme des jeunes ;
- lorsque l'on dispose d'un système d'informations permettant d'obtenir des données enregistrées systématiquement depuis de nombreuses années. Pour réaliser une telle étude, il est nécessaire de disposer de multiples points avant l'intervention, pour évaluer la tendance préexistante et les variations saisonnières éventuelles. Il est également nécessaire de disposer de multiples points après l'intervention, pour déterminer l'effet de l'intervention en prenant en compte la tendance sous-jacente. De nombreuses techniques d'analyse statistique sont disponibles, selon qu'il y ait ou non autocorrélation entre les différents points (selon que les points de mesure effectués à des périodes rapprochées ont plus tendance à être similaires que des points de mesure collectés à des dates éloignées – par exemple, les durées de séjours hospitaliers recueillies durant l'été ont plus tendance à être similaires que les durées de séjours recueillies en juillet, comparées à celles du mois de décembre), en particulier des méthodes utilisant des moyennes mobiles autocorrélées (modèle ARIMA).

Un exemple de séries chronologiques est illustré dans l'encadré n°4.

#### **Encadré n°4 – Exemple de série chronologique**

Ricordeau P, Durieux P, Weill A, Chatellier G, Vallier N, Bissery A, et al. Effect of a nationwide program of educational outreach visits to improve the processes of care for patients with type 2 diabetes. *Int J Technol Assess Health Care* 2003;19(4):705-10.(17).

**Objectif :** évaluer la mise en place d'un programme de « visites à domicile », auprès des médecins généralistes français, destinées à améliorer la prise en charge du diabète de type 2.

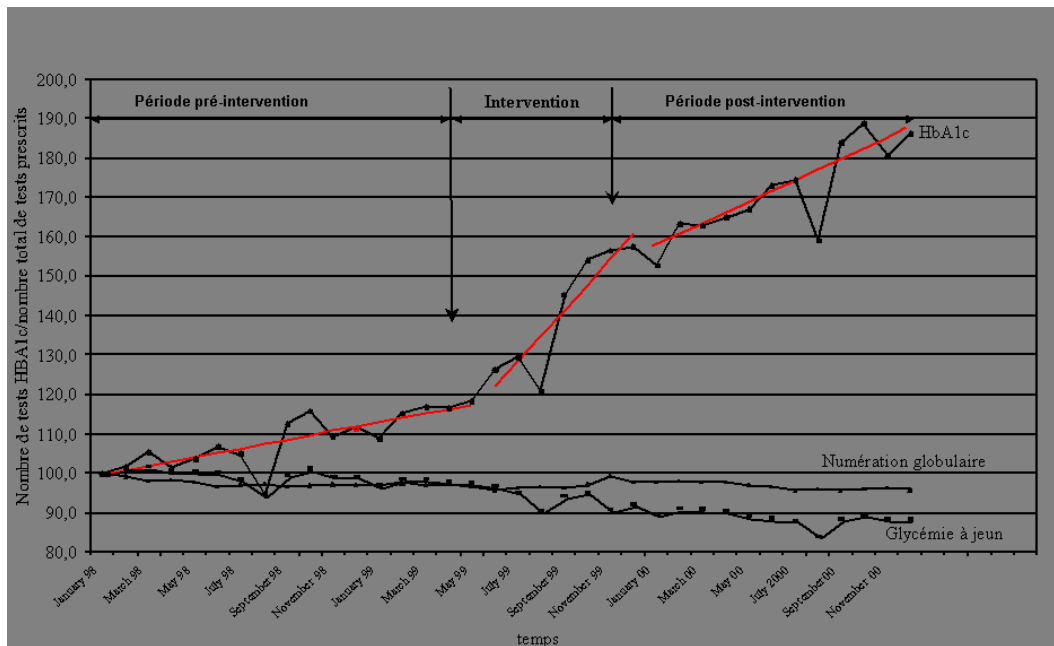
**Plan expérimental :** suivie en série chronologique de données de prescription issues des bases de données de l'Assurance maladie, 17 mois avant, 6 mois pendant et 12 mois après intervention.

**Participants :** Ensemble des médecins français ayant effectué une demande d'ALD, pour un de leur patient durant la période de l'intervention.

**Interventions :** visites à domicile ou entretien téléphonique effectués par un médecin conseil auprès des médecins prescripteurs d'ALD, afin de leur diffuser les recommandations nationales concernant le suivi du diabète de type 2 (notamment la nécessité de doser l'HbA1c tous les 4 mois).

**Critère de jugement :** nombre mensuel de prescriptions d'HbA1c rapporté au nombre total de prescriptions d'examen de laboratoire.

**Résultats :** la rupture de pente observée durant la période de l'intervention montre son efficacité.



## **Pour en savoir plus**

### **Exemples pratiques d'études de type chronologique**

Black N, Hutchings A. Reduction in the use of surgery for glue ear : did national guidelines have an impact? *Qual Saf Health Care* 2002;11(2):121-4.

Bloor K, Freemantle N, Khadjesari Z, Maynard A. Impact of NICE guidance on laparoscopic surgery for inguinal hernias : analysis of interrupted time series. *BMJ* 2003;326(7389):578.

Foy R, Ramsay CR, Grimshaw JM, Penney GC, Vale L, Thomson A, *et al.* The impact of guidelines on mild hypertension in pregnancy : time series analysis. *BJOG* 2004;111(8):765-70.

Mol PG, Wieringa JE, Nannanpanday PV, Gans RO, Degener JE, Laseur M, *et al.* Improving compliance with hospital antibiotic guidelines : a time-series intervention analysis. *J Antimicrob Chemother* 2005;55(4):550-7.

Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, *et al.* What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *BMJ* 2004;329(7473):999.

### **Quelques références méthodologiques**

Ansari F, Gray K, Nathwani D, Phillips G, Ogston S, Ramsay C, *et al.* Outcomes of an intervention to improve hospital antibiotic prescribing : interrupted time series with segmented regression analysis. *J Antimicrob Chemother* 2003;52(5):842-8.

Ostrom CW. *Time series analysis : regression techniques.* London:Sage, 1990.

Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE. Interrupted time series designs in health technology assessment : lessons from two systematic reviews of behavior change strategies. *Int J Technol Assess Health Care* 2003;19(4):613-23.

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther* 2002;27(4):299-309.

## **2.2 Séries en périodes alternées**

Dans certaines situations, il peut être intéressant de réaliser une série chronologique en périodes alternées (*alternate-month design*), c'est-à-dire des alternances de périodes d'une durée plus ou moins courte (par exemple de 1 mois), avec et sans intervention correspondant à une répétition d'études avant/après. Ce type de schéma expérimental ne peut être utilisé qu'en l'absence d'effet rémanent de l'intervention lors des périodes sans intervention.

Ce plan expérimental n'a pas les inconvénients d'une simple étude avant/après, puisque les interventions sont répétées dans le temps. Il permet de plus d'évaluer l'effet d'apprentissage au cours du temps.

Ce plan peut être particulièrement intéressant lorsque l'on souhaite évaluer l'impact d'une intervention dans une seule structure de soins, par exemple un hôpital rendant très difficile la réalisation d'un essai contrôlé randomisé. Il impose cependant que l'événement à mesurer soit suffisamment fréquent pour qu'un nombre suffisant soit recueilli au cours de chaque période.

Un exemple de ce type d'étude est présenté dans l'encadré n°5.

### **Encadré n°5 – Exemple d'étude avec série de périodes alternées**

*Durieux P, Nizard R, Ravaud P, Mounier N, Lepage E. A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior. JAMA 2000;283(21):2816-21. (18)*

**Objectif :** évaluer l'impact d'un système d'aide à la décision informatisé, pour la prévention des thromboses veineuses dans un service de chirurgie orthopédique.

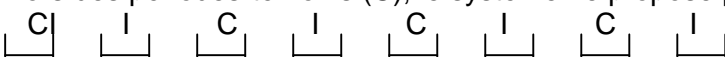
**Plan expérimental :** Série chronologique d'une durée de 18 mois comprenant 4 périodes d'intervention de 10 semaines (système d'aide activé) alternant avec 4 périodes témoins (système d'aide désactivé) de 10 semaines

**Participants :** tous les chirurgiens du service de chirurgie orthopédique

**Interventions :** recommandations pour la prévention des thromboses veineuses informatisées. Système critique, c'est-à-dire intervenant une fois la prescription effectuée, pour proposer une modification si la prescription est contraire aux recommandations.

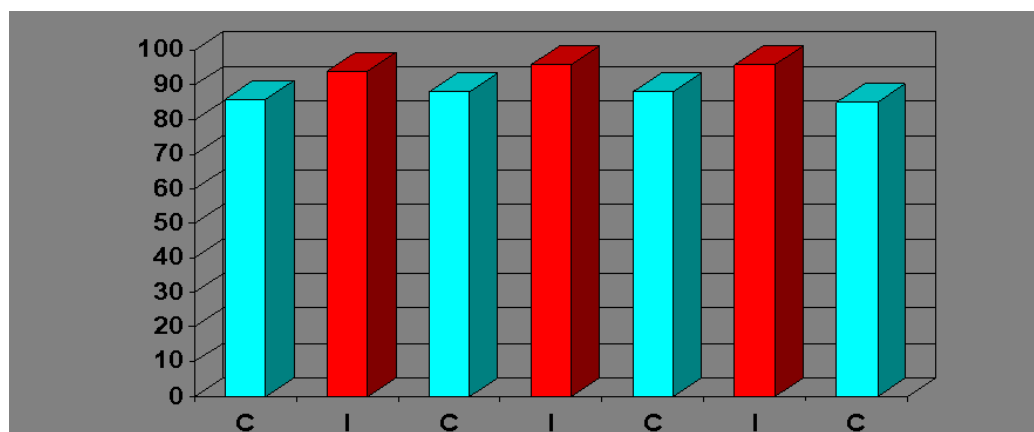
Lors des périodes d'intervention (I), le système propose une prescription adaptée aux facteurs de risque du malade.

Lors des périodes-témoins (C), le système ne propose pas de prescription.



**Critère de jugement :** taux d'adéquation des recommandations, données recueillies automatiquement par le système d'informations.

**Résultats :** l'adéquation aux recommandations augmente lors de chaque période d'intervention, et baisse chaque fois que l'on stoppe l'aide. Il n'y a pas d'effet éducationnel.



**Adéquation aux recommandations :**

périodes témoins : 82,8% [77,6% - 87,1 %]

périodes interventions : 94,9% [92,5% - 96,6%]

## **Pour en savoir plus**

### **Exemples pratiques d'études de type series en périodes alternées**

Colombet I, Bura-Riviere A, Chatila R, Chatellier G, Durieux P. Personalized versus non-personalized computerized decision support system to increase therapeutic quality control of oral anticoagulant therapy : an alternating time series analysis. *BMC Health Serv Res* 2004;4(1):27.

Hay JA, Maldonado L, Weingarten SR, Ellrodt AG. Prospective evaluation of a clinical guideline recommending hospital length of stay in upper gastrointestinal tract hemorrhage. *JAMA* 1997;278(24):2151-6.

Weingarten SR, Riedinger MS, Conner L, Lee TH, Hoffman I, Johnson B, *et al.* Practice guidelines and reminders to reduce duration of hospital stay for patients with chest pain. An interventional trial. *Ann Intern Med* 1994;120(4):257-63.

## **2.3 Études avant/après avec site contrôle contemporain de l'intervention**

Dans ce type d'étude, on cherche à trouver une population-témoins ayant les mêmes caractéristiques que la population qui va subir l'intervention, et l'on recueille les données avant et après intervention dans les deux populations étudiées. La difficulté majeure de ces études est de trouver une population-témoins comparable en tout point à la population-intervention. Si les sites témoins sont étudiés de manière contemporaine aux sites interventions, ce plan expérimental permet d'éviter certains problèmes classiques inhérents aux études avant/après simples.

La difficulté de ce type d'étude est le risque de non-comparabilité avant l'intervention entre les sites-témoins et les sites-interventions. Si cette non-comparabilité existe, l'analyse ultérieure de l'étude est compromise. Parfois dans cette situation, des analyses intragroupes sont réalisées dans chacun des deux groupes (sites-témoins d'une part, sites-intervention d'autre part), en prenant pour hypothèse que s'il existe une variation dans le groupe-intervention et pas de variation dans le groupe-témoins, alors l'intervention est efficace. En réalité, ce type d'analyse est inapproprié en particulier, car si les deux groupes ne sont pas initialement comparables, on peut aussi craindre que la tendance séculaire ne soit pas identique dans les deux groupes, rendant l'interprétation des variations observées erronées.

Un exemple de ce type d'étude est présenté dans l'encadré n°6.

## **Encadré n°6 – Exemple d'étude avant/après avec groupe-témoins**

*Ray-Coquard I, Philip T, de Laroche G, Froger X, Suchaud JP, Voloch A, et al. A controlled "before-after" study : impact of a clinical guidelines programme and regional cancer network organization on medical practice. Br J Cancer 2002;86(3):313-21.(19).*

*Objectif* : évaluer l'impact de l'introduction de recommandations, sur la prise en charge du cancer du sein et du cancer du colon dans un réseau d'hôpitaux anticancéreux.

*Plan expérimental* : étude avant/après contrôlée. Mise en œuvre de recommandations dans les hôpitaux de la région Rhône-Alpes, une région comparable servant de groupe-témoins.

*Participants* : 15 hôpitaux publics et 11 structures privées (région-intervention).

*Interventions* : élaboration puis diffusion de recommandations au sein du réseau d'hôpitaux et structures privées de la région-intervention. Aucune intervention spécifique dans la région-témoin.

*Critère du jugement* : pourcentage d'adéquation aux recommandations, par analyse d'un échantillon de dossiers avant et après intervention, dans la région-intervention et la région-témoin.

*Résultat* : dans le groupe-intervention, l'adéquation aux recommandations passe de 12 % à 36 % (cancer du sein), et de 14 % à 46 % (colon) ; dans le groupe-témoins, les pourcentages passent respectivement de 6 % à 7 %, et de 32 à 39 %.

## **Pour en savoir plus**

### **Exemples pratiques d'études Avant/Après**

Etter JF, Laszlo E. Evaluation of a poster campaign against passive smoking for World No-Tobacco Day. *Patient Educ Couns* 2005;57(2):190-8.

Katz DA, Muehlenbruch DR, Brown RB, Fiore MC, Baker TB, AHRQ Smoking Cessation Guideline Study Group. Effectiveness of a clinic-based strategy for implementing the AHRQ Smoking Cessation Guideline in primary care. *Prev Med* 2002;35(3):293-301.

Leape LL, Cullen DJ, Clapp MD, Burdick E, Demonaco HJ, Erickson JI, *et al.* Pharmacist participation on physician rounds and adverse drug events in the intensive care unit. *JAMA* 1999;282(3):267-70.

Martens JD, Winkens RA, van der WT, de BD, Severens JL. Does a joint development and dissemination of multidisciplinary guidelines improve prescribing behaviour : a pre/post study with concurrent control group and a randomised trial. *BMC Health Serv Res* 2006;6:145.

Reed RL, Revel AD, Carter AO, Saadi HF, Dunn EV. A controlled before-after trial of structured diabetes care in primary health centres in a newly developed country. *Int J Qual Health Care* 2005;17(4):281-6.

### **Quelques références méthodologiques**

Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003;12(1):47-52.

Fergusson D, Hebert P, Shapiro S. The before/after study design in transfusion medicine : methodologic considerations. *Transfus Med Rev* 2002;16(4):296-303.

Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference.* Boston:Houghton Mifflin Company, 2002.

## 3. Autres particularités méthodologiques des études d'intervention

### 3.1 Standardisation de l'intervention

Pour toute étude visant à évaluer l'impact d'une intervention, il est impératif de standardiser l'intervention effectuée, afin que d'autres puissent reproduire ultérieurement cette intervention si celle-ci s'avère efficace. Il est donc impératif d'être extrêmement précis sur la recommandation utilisée (tout ou partie de la recommandation), la méthode de mise en œuvre, et la population concernée.

Il est ainsi important de préciser la recommandation de pratique clinique, sur laquelle on s'appuie pour réaliser l'intervention (recommandation issue de la HAS, d'une société savante ou recommandation issue d'une agence étrangère ou d'une société savante étrangère et traduite), et même les points particuliers de la recommandation qui sont les plus importants, et qui pourront constituer les critères de jugement à propos desquels on jugera de l'efficacité de l'intervention. Il est également primordial de définir les méthodes de dissémination et de mise en œuvre qui vont être utilisées. Sur ce point, il faut bien-sûr prendre en compte le fait que de nombreuses méthodes ont déjà largement été évaluées, et ont fait la preuve de leur efficacité ou au contraire de leur inefficacité (2).

Enfin, la population-cible (médecins spécialistes ou généralistes, hospitaliers ou non hospitaliers, hôpital universitaire ou général, etc.) de cette intervention doit être parfaitement définie, car une intervention peut être efficace dans un contexte de soins, et totalement inefficace dans un autre contexte de soins.

### 3.2 Critères de jugement ou indicateurs

#### 3.2.1 Évaluer les connaissances ou évaluer les pratiques réelles

Les critères de jugement définis pour chaque type d'étude ne doivent pas être des critères visant à évaluer l'accord des médecins, avec les recommandations ou les connaissances des médecins ou même des pratiques déclarées par les médecins (2). Les critères de jugement sont destinés à mesurer l'impact, positif et négatif, d'une intervention, en termes de modification réelle des pratiques des médecins, et non de modifications de leurs connaissances ou de leurs pratiques déclarées.

Cependant, l'utilisation de *case-vignettes* sous des formes diverses (papier ou informatique) a été proposée et utilisée depuis de nombreuses années. Ces cas-papiers ou électroniques sont des brèves histoires d'un malade fictif basé sur une situation clinique réaliste, et accompagné par une ou plusieurs questions qui explorent ce qu'un médecin ferait si un tel malade se présentait.

Cette méthode a été validée par comparaison à d'autres méthodes d'évaluation de la qualité des soins : malades simulés (*standardized patients*) et audit de dossiers. Cette méthode ne semble pas, à ce jour, avoir été utilisée dans le cadre d'études d'intervention, mais elle pourrait l'être, en particulier dans les situations où des différences de recrutement entre hôpitaux ou structure de soins constituent un problème. En effet, elle permet d'avoir des réponses attribuables au seul médecin, et indépendantes de la variabilité liée au recrutement (20-24).

#### 3.2.2 Évaluer les procédures ou les résultats

Les critères habituellement utilisés sont des critères évaluant les procédures, plutôt que des critères évaluant les résultats pour plusieurs raisons. La première est que l'objectif de l'intervention est d'améliorer les pratiques professionnelles en faisant appliquer des recommandations de pratique clinique, il semble donc logique d'utiliser des critères évaluant



ces pratiques professionnelles. Ainsi, si une recommandation indique qu'il faut réaliser une coloscopie tous les deux ans sur une population à risque, le critère de jugement sera le pourcentage de patients ayant une coloscopie selon la recommandation. La seconde est que les recommandations sont basées sur des résultats d'études publiées, et fréquemment sur les résultats des essais randomisés qui constituent le meilleur niveau de preuve. Si les essais thérapeutiques démontrent par exemple qu'il est utile de prescrire de l'aspirine en prévention secondaire après un infarctus du myocarde, il est inutile, si l'on veut évaluer l'impact de recommandations visant à améliorer la qualité de la prescription après infarctus du myocarde, de vouloir utiliser un critère de jugement de type résultat (par exemple la mortalité après infarctus du myocarde). C'est inutile, car c'est l'objet des essais randomisés ayant conduit à la recommandation, et non d'un essai ayant pour objet d'évaluer l'impact de la mise en œuvre de recommandations ou pour objet de démontrer l'impact sur la mortalité d'une telle prescription. Une telle ambition serait en effet irréaliste, car s'il faut des milliers ou des dizaines de milliers de malades pour démontrer, dans le cadre d'un essai thérapeutique, l'intérêt de l'aspirine, il faudra un nombre infiniment supérieur de malades pour démontrer l'impact d'une intervention qui va permettre d'augmenter de 60 à 80 % le pourcentage de malades qui recevront réellement de l'aspirine après un infarctus du myocarde. Mant et Hicks (25) ont montré que pour mettre en évidence une réduction relative de mortalité de 3 % entre deux hôpitaux, il était nécessaire d'inclure 32 846 malades par hôpital, pour un critère de jugement de type résultat (mortalité), et 155 si l'on utilise un autre critère de jugement de type procédure de soins (Est-ce que les malades devant recevoir de l'aspirine ou des bêtabloquants recevaient bien ces traitements ?).

### 3.3 Aveugle

Les critères de jugement ou indicateurs retenus doivent idéalement être recueillis à l'insu de l'intervention réalisée, afin d'éviter un biais de performance. Ce type de biais n'existe pas, bien évidemment si les données utilisées sont recueillies de manière systématique, indépendamment de l'étude. *A contrario* si ces données sont recueillies spécifiquement pour une étude, il est logique de prévoir un évaluateur (*outcome assessor*) identique pour les différents bras ou différents sites, n'ayant pas été impliqué dans la réalisation des interventions, et idéalement n'ayant pas connaissance de l'intervention effectuée lorsqu'il réalise cette évaluation (cette condition n'est par exemple pas respectée dans la *référence 16* citée plus haut).

### 3.4 Sources de données

Lorsque l'on désire évaluer une intervention, on peut utiliser soit des données recueillies spécifiquement pour l'étude, soit des données recueillies de manière systématique, indépendamment de l'étude. Les données recueillies spécifiquement pour l'étude ont un avantage majeur, qui est que l'on peut décider parfaitement quels sont indicateurs les plus pertinents en fonction de l'intervention étudiée, et des pratiques que l'on veut modifier. On peut également recueillir des données plus fines, qui nous permettent par exemple de tenter de comprendre pourquoi une recommandation n'est pas appliquée, et les causes de cette non-application.

Ces données seront recueillies en utilisant les méthodes habituelles d'audit de pratiques cliniques, de revues de dossiers, etc.

Dans certaines circonstances, par exemple pour évaluer les pratiques infirmières, l'évaluation des pratiques est plus complexe, et le risque d'une modification des pratiques en présence d'un observateur indépendant ayant pour objet d'évaluer ces pratiques ne doit pas être négligé.

Le recueil de données spécifiques a souvent un coût élevé, et nécessite du personnel entraîné et dédié.

L'utilisation de données recueillies de manière systématique, indépendante de l'intervention est une autre option. Dans de nombreux domaines, prescription des examens complémentaires ou prescription des médicaments, il existe de multiples systèmes d'informations, par exemple à l'hôpital, mais également en médecine de ville (données de l'Assurance maladie). Bien évidemment, les informations disponibles à l'aide de ces systèmes d'informations sont souvent incomplètes, et parfois inexactes (par exemple pour le diagnostic des malades) (26), car ces systèmes d'informations n'ont pas été conçus dans cet objectif. Néanmoins, les sources d'informations ont l'avantage majeur d'exister (parfois depuis plusieurs années), et de ne pas nécessiter de moyens supplémentaires pour le recueil des données.

# Concevoir et mettre en œuvre une étude

Ce chapitre présente les différentes étapes indispensables à la conception et à la réalisation d'une étude destinée à évaluer, selon un schéma valide, l'impact d'une intervention sur la qualité des soins. Les étapes présentées concernent principalement la réalisation d'une étude d'intervention contrôlée randomisée. Cependant, la plupart des chapitres s'appliquent aussi à d'autres types d'études, étude avant/après, avec groupe-témoin ou série chronologique.

Nous distinguerons trois périodes chronologiques distinctes, avant, pendant et après la réalisation de l'étude.

## 1. Avant la réalisation de l'étude

Avant la réalisation d'une étude d'intervention, il est important de s'assurer de l'intérêt et de l'implication des parties prenantes. Il est aussi essentiel d'identifier toutes les structures et ressources locales, régionales ou nationales le plus en amont possible.

Le processus de mesure et, en fin de compte, l'amélioration de la qualité des soins impliquent une grande variété d'intervenants : des professionnels travaillant au sein de la même équipe, des professionnels provenant de différentes équipes ou spécialités, des personnes en charge de l'amélioration de la qualité, des méthodologistes, des représentants de structures administratives, des patients ou des spécialistes en sciences sociales. Le fait d'identifier, et d'impliquer dès le départ toutes les parties prenantes, permet de garantir le traitement des problèmes susceptibles de survenir, et d'obtenir un soutien efficace à la réalisation de l'étude.

Enfin, il faut réfléchir aux questions éthiques. Les exigences relatives au consentement éclairé et à la confidentialité varient selon le type d'étude, le type d'intervention (selon que cette intervention s'adresse aux professionnels de santé ou aux patients) et le type de données recueillies. Il est donc nécessaire de se renseigner auprès des structures compétentes (cf. plus loin).

Il peut être utile à cette phase de créer un groupe de travail comprenant toutes les parties prenantes au projet d'étude, qui sera coordonné par un responsable de projet.

La mise en place d'une étude d'intervention requiert la succession de six étapes listées ci-dessous :

- La formulation du problème :
  - Choix du thème de travail, formulation du problème, analyse des pratiques existantes.
- La détermination de la population-cible :
  - Détermination du groupe-cible de l'intervention, du lieu de l'étude et des objectifs de changement.
- Le choix de la stratégie d'amélioration :
  - Choix d'une stratégie d'intervention adaptée, prise en compte des freins et facilitations potentielles à sa mise en œuvre, élaboration des éléments précis et détaillés de cette stratégie.
- Le choix du plan de l'étude :
  - Choix du plan d'étude expérimental ou d'observation.
- Le choix et recueil des indicateurs permettant de juger l'impact de l'intervention.
- La détermination des moyens humains et matériels nécessaires à sa mise en œuvre.

Ces six étapes sont synthétisées dans la suite de ce document, dans lequel sont également précisés les différents éléments qui doivent figurer dans un protocole.

**La méthode PICOT (*Population – Intervention – Comparison – Outcomes – Time*) peut aider à la mise en place de ces différentes étapes.**

<b>P</b>	Population – Trois niveaux de population peuvent être définis :
	(P1) Patients = patients concernés par l'intervention (par exemple, des patients atteints de diabète de type 2)
	(P2) Professionnels = professionnels de santé visés par l'intervention (par exemple, des médecins généralistes)
	(P3) Services médicaux = groupes de professionnels visés par l'intervention (par exemple, cabinets de médecine générale) ou l'environnement dans lequel les professionnels concernés travaillent si l'intervention vise des individus plutôt que des groupes de professionnels. (par exemple, cabinets de médecine générale en zone rurale)
<b>I</b>	Intervention – Quelle intervention évaluer ? (par exemple, intervention faisant appel à la formation médicale continue)
<b>C</b>	Comparaison – Quelle intervention utiliser pour la comparaison ? (intervention alternative ou absence d'intervention) (par exemple, formation interne habituelle)
<b>O</b>	<i>Outcome</i> ou Résultats – Comment l'effet de l'intervention sera-t-il mesuré ? (par exemple, taux de respect d'une recommandation, pourcentage de patients diabétiques ayant eu un dosage d'hémoglobine glyquée dans les derniers six mois)
<b>T</b>	Temps – Combien de temps attendre après l'arrêt de l'intervention avant d'évaluer les résultats ? (par exemple, un an après la fin de l'intervention)

## 1.1 La formulation du problème

### 1.1.1 Thème de travail

Il paraît évident que le choix du sujet est un important critère de réussite. Le sujet doit concerner un problème clinique important en termes de fréquence, gravité ou risque pour le patient (ou les professionnels de santé). Il doit exister des éléments qui font penser à une pratique non adéquate, et le besoin d'amélioration doit être ressenti par les professionnels concernés. Enfin, on doit disposer de tous les éléments permettant l'amélioration des pratiques : référentiel existant valide ou d'élaboration facile, stratégie d'amélioration facile à mettre en place, capacité à recueillir un indicateur d'évaluation valide. Il est en effet indispensable d'envisager tous les aspects de la démarche d'amélioration, dès le choix du thème. Une grille de choix est présentée en annexe.

#### **Protocole :**

Donner des informations importantes du contexte expliquant pourquoi la maladie ou l'état constitue un problème, et pourquoi il est nécessaire d'améliorer la qualité des soins.

Prévoir une analyse détaillée des études antérieures pertinentes concernant le thème clinique, le référentiel éventuel et l'intervention envisagée.

Indiquer les sources bibliographiques prises en compte.

### 1.1.2 La justification de la recherche

La Déclaration d'Helsinki stipule que la recherche biomédicale, impliquant des humains doit être fondée sur une connaissance approfondie de la littérature scientifique – c'est-à-dire que le fait d'exposer inutilement des êtres humains aux risques de la recherche est contraire à l'éthique. Certaines études cliniques se sont avérées inutiles, parce que la réponse à la question posée aurait été (ou aurait pu être) obtenue par le biais d'une revue systématique de la littérature existante.

Il est souhaitable que la justification de l'étude décrive les raisons pour lesquelles la question posée est importante. La justification doit inclure les bienfaits perceptibles de toute intervention devant être évaluée. Elle doit indiquer la manière dont l'intervention étudiée pourrait fonctionner, notamment si les données antérieures concernant ce type d'intervention sont peu nombreuses, voire inexistantes.

#### **Protocole :**

Décrire la raison pour laquelle la question de recherche mérite d'être étudiée.

Décrire comment l'intervention évaluée pourrait fonctionner.

Préciser comment les résultats pourraient faire progresser les connaissances théoriques et pratiques, dans le domaine concerné.

### 1.1.3 Existence de recommandations

Si l'intervention concerne la mise en œuvre de recommandations destinées à améliorer la pratique clinique, il est important d'utiliser un référentiel valide, et qui sera accepté par les utilisateurs. Ce référentiel doit être issu de recommandations élaborées, selon un processus rigoureux. Il existe aujourd'hui suffisamment de recommandations nationales publiées pour ne pas avoir à élaborer des recommandations locales. En France, ces recommandations sont disponibles sur le site de la HAS ou pour la cancérologie sur le site de la Fédération nationale des centres de lutte contre le cancer ([www.fnclcc.fr](http://www.fnclcc.fr)). On peut aussi prendre en compte une recommandation publiée par un organisme étranger comme le *National Institute for Health and Clinical Excellence* (NICE) au Royaume-Uni ([www.nice.org.uk](http://www.nice.org.uk)) ou une recommandation référencée par la *National Guideline Clearinghouse™* (NGC) de l'*Agency for Healthcare Research and Quality* (AHRQ) aux États-Unis ([www.guideline.gov](http://www.guideline.gov)).

Il faut cependant dans tous les cas s'assurer de la qualité méthodologique de la recommandation, et prendre garde à ce que son contenu ne soit pas en contradiction avec la réglementation en vigueur en France (en particulier avec les données de l'Autorisation de mise sur le marché en ce qui concerne les médicaments). La qualité d'une recommandation peut être évaluée au moyen d'une grille de lecture comme celle proposée par le groupe AGREE (*Appraisal of Guidelines Research and Evaluation*) ([www.agreecollaboration.org](http://www.agreecollaboration.org)), qui a fait l'objet d'une traduction en français. Il est en particulier essentiel de s'assurer que l'on pourra extraire de la recommandation des critères ou indicateurs d'évaluation simples.

Si une situation clinique d'intérêt n'a pas fait l'objet de recommandations récentes ou si les pratiques idéales sont discutées, il faut envisager de changer de thème de travail.

### 1.1.4 Formulation de la question et des objectifs de recherche

Le but d'une étude d'intervention est d'évaluer si (et à quel degré) une ou plusieurs interventions mises en œuvre au niveau d'une population donnent ou non des résultats différents (ex. : santé, processus de soins de santé, économie, satisfaction des patients) par rapport à une ou plusieurs autres interventions ou à l'absence d'intervention. Il est souhaitable que le « but de l'étude » décrive, en une ou deux phrases, la finalité de l'étude proposée.

Les objectifs sont les questions spécifiques pour lesquelles l'étude est destinée. Il y a généralement une question principale et éventuellement plusieurs questions secondaires.

**Protocole :**

Énoncer la question de recherche.

Spécifier les hypothèses à explorer et les objectifs de recherche.

S'assurer que la question inclut : la population, l'intervention expérimentale, l'intervention de comparaison, les résultats clés et le Temps = « PICOT »

Établir une liste claire et concise énumérant les objectifs principaux et secondaires, ainsi que les hypothèses que l'étude va examiner.

Ceux-ci doivent être traités en termes de résultats mesurables.

### 1.1.5 Évaluation des pratiques avant intervention

La mesure de la performance de base est indispensable à la planification et à l'évaluation d'une intervention, car elle aide à :

- justifier l'intervention en démontrant l'existence d'un écart entre la pratique clinique réelle et la pratique clinique souhaitée ;
- préciser quelle sera la population-cible de l'étude
- évaluer l'ampleur du problème. Une faible performance de base indique une importante marge potentielle d'amélioration, tandis qu'une performance de base élevée indique une petite marge pour l'amélioration (effet plafond) ;
- identifier les modes de pratique et les facteurs qui les déterminent. On peut alors utiliser ces facteurs pour adapter l'intervention ;
- utiliser les mesures comme partie intégrante d'une intervention impliquant un retour d'informations ;
- calculer la taille de l'échantillon.

Les moyens étant limités, il est important de ne pas mettre en place une intervention destinée à améliorer des pratiques déjà satisfaisantes.

L'analyse des pratiques peut être réalisée de différentes manières : analyse rétrospective de dossiers médicaux, analyse d'indicateurs issus du système d'informations, recueil prospectif de données sur une période limitée, voire s'il n'est pas possible de faire autrement, enquêtes auprès de professionnels de santé sur leurs pratiques déclarées ou sur leurs connaissances concernant une pratique donnée. Les pratiques déclarées ne sont pas un bon reflet des pratiques réelles (2,27).

L'analyse des dossiers dans le cadre d'un audit clinique a fait l'objet de plusieurs documents méthodologiques disponibles sur le site de la HAS. Rappelons simplement que l'intérêt de cette méthode est limité par le temps nécessaire à la recherche et l'analyse du contenu des dossiers médicaux ainsi que la qualité des informations retrouvées dans les dossiers.

Recueillir des données de façon prospective, que l'on recherche des informations sur les pratiques réelles, déclarées ou les connaissances, permet de s'assurer de la qualité du recueil des informations, mais demande également du temps et des moyens humains.

L'analyse des indicateurs issus du système d'informations est limitée par le type d'informations que l'on peut obtenir (essentiellement des données de prescription d'actes ou de médicaments), et par l'accès à un système adéquat d'informations. Cette approche doit être privilégiée chaque fois que possible. Elle est plus simple à mettre en œuvre en milieu hospitalier. En médecine ambulatoire, elle demande l'accès aux systèmes d'informations de l'Assurance maladie.

L'analyse des pratiques implique, dès ce stade, de se demander quels sont les critères de jugement de la qualité des pratiques l'on souhaite mesurer. Ces critères peuvent toucher aux structures, aux procédures ou aux résultats. Pour des raisons développées plus loin, les critères de procédures sont les plus fréquemment utilisés.

Par exemple, si l'on souhaite évaluer la prise en charge du diabète de type 2, l'analyse de la base de données de l'Assurance maladie, où sont enregistrées toutes les prestations remboursées aux assurés sociaux, permet d'estimer le pourcentage de patients chez qui une recherche d'hémoglobine glyquée a été réalisée dans les 6 derniers mois (17). Un recueil prospectif spécifique permet de recueillir non seulement le nombre d'exams réalisés, mais aussi leur résultat, ainsi qu'un certain nombre d'informations impossibles à recueillir par l'analyse des bases de données de l'Assurance maladie (réalisation d'un fonds d'œil, mesure de la pression artérielle, etc.) (28). Ces dernières informations auraient probablement pu aussi être recueillies de façon rétrospective par l'analyse de dossiers médicaux, mais la qualité des informations obtenues aurait certainement été bien inférieure.

## 1.2 Détermination de la population-cible

### 1.2.1 Choix de la population et du site concernés par l'étude

La population étudiée n'est pas le groupe d'individus ou de grappes inclus dans l'étude (à savoir : l'échantillon), mais plutôt le groupe auquel l'on souhaite généraliser les résultats de l'étude. Il faut d'abord décider et définir la population, puis élaborer un plan pour obtenir un échantillon représentatif de cette population.

La méthode PICOT présentée plus haut peut aider à la définition de la population. Il faut également définir l'environnement et le lieu de l'étude, car ils influenceront sur la validité externe des résultats. Les différentes institutions de santé varient grandement en termes d'organisation, d'expérience, de ressources et ainsi qu'en termes de risque de référence de l'état pathologique faisant l'objet de l'étude. Le climat et d'autres facteurs physiques, économiques, géographiques et socioculturels peuvent également affecter la validité externe d'une étude. En outre, il convient de spécifier clairement si l'étude doit être effectuée dans un ou plusieurs centres.

#### **Protocole :**

Fournir une description de la population de l'étude en termes de patients, professionnels et services médicaux.

Donner la raison pour laquelle ces patients, professionnels et services médicaux ont été sélectionnés.

Décrire et expliquer l'environnement et les lieux où l'étude sera effectuée.

Indiquer si l'étude a eu lieu dans un ou plusieurs centres (études multicentrique).

### 1.2.2 Critères d'inclusion et de non-inclusion

Il est généralement nécessaire de restreindre la population étudiée au moyen de critères d'inclusion. Il existe trois niveaux de participants dans une étude d'intervention en grappes : patients, professionnels et services. Il doit y avoir des critères explicites d'inclusion pour chacun d'entre eux.

En ce qui concerne les patients, les critères d'inclusion sont généralement liés à leur maladie ou à leur état (par exemple : fibrillation auriculaire, femmes en phase d'accouchement). Des critères supplémentaires peuvent inclure des facteurs sociodémographiques (âge, sexe, etc.), un facteur de comorbidité (ex : diabète) ou un type d'intervention reçue (ex : examen, intervention chirurgicale, traitement médicamenteux) ; les critères de non-inclusion sont souvent utilisés pour garantir la sécurité des patients.

En ce qui concerne les professionnels de santé, les critères d'inclusion peuvent inclure la profession (infirmière, médecin), la spécialité clinique (infirmière en traumatologie, cardiologue), le nombre minimum de patients pris en charge, l'ancienneté (étudiant, praticien diplômé), etc.

En ce qui concerne les services médicaux, les critères d'inclusion peuvent inclure le niveau des soins prodigués (secondaires, tertiaires), le nombre minimum de cas pris en charge, etc.

Une description précise de ces critères, y compris les critères démographiques de la population étudiée, est nécessaire pour permettre aux lecteurs d'évaluer la validité externe (l'aptitude à la généralisation) des résultats de l'étude.

**Protocole :**

Établir des listes séparées de critères d'inclusion pour les patients, les professionnels et les services médicaux.

Établir des listes séparées de critères d'inclusion et de non-inclusion.

### 1.3 Méthodes de recrutement

Le recrutement d'un nombre adéquat de participants (patients, professionnels de santé et services médicaux) constitue un critère essentiel de succès. Chaque essai en grappes est confronté au défi d'atteindre ses objectifs de recrutement dans des délais limités, et avec un budget restreint. Certaines études y parviennent, mais beaucoup d'autres sont contraintes de prolonger la période de recrutement et/ou d'accroître les efforts et les coûts. Les principaux problèmes de recrutement sont souvent liés à une mauvaise planification à tous les niveaux de l'étude, une surestimation du recrutement à partir d'une source spécifique, et l'incapacité à mettre rapidement en œuvre des stratégies de recrutement, si celui-ci se déroule plus lentement que prévu.

La méthode de recrutement des participants est très importante, car elle peut fausser une étude en produisant des écarts qui seront impossibles à corriger par la suite. Il faut bien réfléchir à cette procédure de sélection, afin de pouvoir situer l'échantillon en termes de représentativité par rapport à la population totale non mesurée. On peut examiner le cas de quelques centres qui refusent de participer ; si leur refus est dû à leur faible disposition au changement, peut-être à cause d'une gestion médiocre, la sélection sera faussée, et l'on ne disposera, dans l'étude, que de centres de grande qualité, tout à fait disposés au changement. Il est presque impossible de connaître la raison d'un refus, dès lors qu'un centre a décliné une invitation à participer à une étude.

**Protocole :**

Décrire la stratégie ou le procédé utilisé pour sélectionner et recruter les participants individuels et les grappes pour l'étude.

### 1.4 Choix de la stratégie d'amélioration

#### 1.4.1 Prise en compte des freins à l'amélioration des pratiques

Le faible impact des recommandations n'est pas uniquement lié à la méthodologie de leur mise en œuvre, mais également par l'existence de facteurs particuliers, qui constituent des freins ou barrières à leur mise en œuvre. Il peut donc être utile d'estimer leur existence et leur importance, avant la mise en œuvre d'une intervention destinée à améliorer la qualité des soins. Le choix de l'intervention peut alors être choisi spécifiquement, de manière à prendre en compte ces barrières ou freins, permettant ainsi un meilleur impact.



Les freins liés à l'application des recommandations peuvent être classés en facteurs liés :

- aux caractéristiques individuelles des professionnels de santé (facteurs cognitifs, comportementaux, motivation) ;
- à l'environnement humain (patients, interaction médecin/patient, interaction entre professionnels) ;
- aux aspects organisationnels (organisation des soins, formation), politiques ou structurels (mode de financement des soins, environnement légal).

Ces freins sont donc susceptibles d'intervenir à chaque phase de la prise en charge d'un malade.

Tout professionnel de santé souhaite délivrer les meilleurs soins possibles à son patient. Mais, comme pour tout comportement (par exemple comportement social, physique, sexuel), il n'y a pas de lien évident entre connaissance et pratique de soins. Le simple lavage des mains dans la prévention des infections nosocomiales en est une bonne illustration (29). Les professionnels de santé connaissent l'importance et la gravité des infections survenant à l'hôpital. Ils savent que le lavage des mains prévient un grand nombre de ces infections. Mais il existe des obstacles individuels (motivation), liés à l'environnement humain (manque de *leadership*), organisationnels (manque de temps) ou financiers (manque de lavabos ou de dispositifs de délivrance de solution hydro-alcoolique) qui empêchent l'amélioration des pratiques. Toute action visant à améliorer les pratiques doit donc prendre en compte l'ensemble de ces obstacles.

De nombreux modèles ou théories sont proposés pour concevoir une stratégie de changement des pratiques médicales. Ces théories sont issues en particulier du domaine de l'apprentissage et de la connaissance, de l'analyse du comportement, de la sociologie du travail, de l'économie et du management, de la qualité. La description de ces théories et de leur utilisation potentielle sort du cadre de ce document.

Un grand nombre de méthodes issues de la recherche qualitative ont été proposées pour identifier les barrières potentielles au changement : questionnaires, *focus groups*, entretiens individuels, méthode Delphi, observation. Les barrières (et facilitations) identifiées, une stratégie de changement adaptée peut être mise en place. Cependant, les données concernant l'efficacité de telles interventions manquent. Comment évaluer et tenir compte des obstacles au changement reste du domaine de la recherche, et demande l'apport de chercheurs en sciences sociales (30-32).

#### 1.4.2 Choix d'une intervention adaptée

Il est important de décrire en détail chaque intervention, y compris l'intervention « témoin ». On peut s'inspirer de la classification proposée par le groupe EPOC (*Effective practice and Organization of Care review group*), qui liste les différents types d'interventions possibles pour modifier le comportement du prescripteur (*annexe 2*). Mais il existe une très grande hétérogénéité dans ces interventions, même si l'on ne prend en compte que celles d'un même type (par exemple l'*audit feedback*), et il n'est pas possible de faire sur ce point des recommandations précises. Toute intervention doit être adaptée au contexte local : moyens disponibles, organisation des soins, existence ou non d'un système d'information, etc. Par exemple, la mise en place de visite à domicile (ou *outreach visits*), telles que les visites confraternelles utilisées par l'Assurance maladie, demande des moyens, et est plus appropriée lorsque l'intervention touche de nombreux centres ou prestataires de soins (comme c'est le cas en médecine ambulatoire). Au contraire, les rappels au moment de la décision, dont l'utilisation est favorisée par l'existence d'un système informatisé de prescription, sont plus adaptés à l'exercice en milieu hospitalier.

La meilleure façon pour choisir une intervention est donc de consulter la banque de données *Cochrane* ([www.cochrane.org](http://www.cochrane.org)). Cette banque contient de nombreuses revues systématiques qui traitent soit d'une intervention spécifique (par exemple l'*audit feedback*), soit d'un problème de santé spécifique (par exemple les interventions destinées à améliorer la prise en charge du

diabète (33)). On peut aussi s'aider du rapport publié en 2000 par l'ANAES, disponible sur le site de la HAS (2). Enfin, il est souvent utile d'extraire des revues systématiques quelques articles clés, afin de voir comment les interventions sont précisément mises en œuvre.

La description détaillée de l'intervention doit comprendre (34) :

- le type de la stratégie d'intervention (classification *Cochrane*),
- le groupe et les participants cibles ;
- le ou les « responsables de la mise en œuvre »/intervenants (experts, professionnels de santé, leaders d'opinion, etc.) ;
- la fréquence, le moment et la durée de l'intervention ;
- la description du type d'information fournie.

Par exemple, pour décrire le responsable de la mise en œuvre d'une intervention pédagogique, il est nécessaire de préciser le nombre, la formation et le niveau de compétence des personnes chargées de la formation.

Cette description sera bien-sûr très utile à l'évaluation de la validité externe des résultats de l'étude.

#### **Protocole :**

Décrire en détail l'intervention ou les interventions qui seront effectuées.

Décrire le déroulement de l'intervention.

Décrire en détail qui se chargera d'effectuer la ou les interventions.

Décrire en détail les « soins courants » qui seront prodigués.

## **1.5 Choix du plan d'étude expérimental ou d'observation**

Pour évaluer une intervention destinée à l'amélioration de la qualité, il est possible de choisir parmi un certain nombre de schémas d'étude, randomisés ou non randomisés, décrits au chapitre 2 :

<b>Schémas randomisés</b>	<b>schémas non randomisés</b>
Essais contrôlés à randomisation individuelle	Études non contrôlées avant et après
Essais randomisés à randomisation collective (en grappes)	Études contrôlées avant et après
	Modèles à série temporelle

Afin de recueillir des informations généralisables sur la valeur d'une intervention expérimentale, on doit être capable d'attribuer avec certitude les effets observés de l'expérimentation à cette intervention. Par conséquent, il est nécessaire de choisir une étude qui minimise les biais éventuels (tendance systématique de tout facteur qui fausse l'interprétation des résultats, et compromet les conclusions tirées de l'étude). Les essais cliniques randomisés par grappes sont considérés comme la conception idéale pour évaluer les interventions d'amélioration de la qualité (cf. chapitre précédent).

Cependant, dans les essais cliniques randomisés par grappes dans lesquels relativement peu de grappes (ex. : hôpitaux) sont disponibles pour la randomisation, il existe un risque de déséquilibre accru dans les résultats observés, dans le groupe expérimental et le groupe-témoin dû au jeu du hasard.

On peut aussi vouloir réaliser une étude dans sa propre structure rendant très difficile la randomisation. Dans tous les cas, si l'on souhaite réellement évaluer l'impact d'une intervention, il faut à tout pris éviter la réalisation d'une étude avant/après, mais envisager une série chronologique ou une étude avant/après avec groupe-témoin (par exemple un service

dans lequel on ne met pas en place l'intervention). Dans ce dernier cas, on doit être attentif au risque de contamination entre les groupes.

Les séries chronologiques (ou temporelles) sont particulièrement appropriées lorsque l'on dispose de bases de données permettant de voir l'évolution d'un indicateur sur une longue période (par exemple bases de données hospitalières ou de l'Assurance maladie).

L'étude avant/après ne permet que de montrer que la qualité des soins s'améliore, sans pouvoir déterminer quelle en est la raison.

Que l'étude soit randomisée ou non, il existe des critères de qualité que l'on doit respecter (cf. *annexe 3*)

#### **Protocole :**

Utiliser si possible un modèle randomisé.

Utiliser un modèle à série temporelle si une conception randomisée n'est pas possible.

Faire attention de ne pas tirer des conclusions sur le résultat d'une étude avant/après.

## **1.6 Choix des indicateurs ou critères de jugement**

### **1.6.1 Indicateurs**

La HAS définit un indicateur comme « une variable qui décrit un élément de situation ou une évolution d'un point de vue quantitatif. » La HAS réserve le terme de critère au cadre de l'audit clinique, et le définit comme « un élément auquel on se réfère pour porter un jugement, une appréciation sur la conformité des pratiques. Chaque critère se caractérise par une variable qui permet la mesure du critère, un standard qui est la valeur attendue, une ou des exceptions le cas échéant et des instructions de recueil. » La HAS réserve le terme d'indicateur aux variables pouvant ultérieurement permettre un suivi de la qualité dans le cadre d'un système de recueil de données (35).

Dans ce texte, nous employons indifféremment les deux termes, dans la mesure où le principe même de l'évaluation des pratiques est d'être de suivre à long terme la qualité, quelle que soit la méthode de mesure entreprise, y compris celle de l'audit clinique.

Il existe deux principaux types de résultats : ceux qui se rapportent aux professionnels de santé (reflétant les procédures de soins) et ceux qui se rapportent aux patients (reflétant les résultats des soins).

*Procédures de soins (professionnels de santé) :*

- adhésion à la recommandation ou à la « meilleure pratique » ;
- indicateurs d'utilisation de ressources (par exemple, nombre de tests prescrits par unité de temps ou rapportés à l'activité).

*Résultat des soins (relatifs aux patients) :*

- mortalité (habituellement provoquée par la maladie ou au trouble visé) ;
- morbidité (généralement liée à la maladie ou à l'état visé) ;
- qualité de vie (générique ou spécifique de la maladie ou de l'état concerné) ;
- résultats de substitution (par exemple, pression artérielle, débit de pointe) ;
- indicateurs d'utilisation de ressources (par exemple : réadmission, durée du séjour).

Il a déjà été précisé précédemment qu'il fallait évaluer les pratiques réelles et non déclarées, et qu'il était en général plus pertinent (et plus simple) d'évaluer les procédures de soins que les résultats.

Par exemple, il a été montré que le fait de prescrire des bêtabloquants après un infarctus du myocarde constituait un facteur préventif de la récurrence. Le pourcentage de patients ayant reçu des bêtabloquants après un infarctus du myocarde est un indicateur de procédure de soins valide et reconnu, qui témoigne de la qualité de la prise en charge de l'infarctus du myocarde.

Il est évidemment plus facile à recueillir que le taux de mortalité par récurrence d'infarctus du myocarde (36).

### 1.6.2 Comment établir un indicateur de procédure ?

Les indicateurs de procédures doivent être issus de recommandations professionnelles valides.

La méthodologie de construction d'un indicateur, critère de jugement de l'impact d'une intervention, a été décrite dans un document disponible sur le site Internet de la HAS (35).

#### **Protocole :**

Établir une liste des indicateurs qui seront utilisés, et justifier leur utilisation.

Identifier le critère de jugement principal.

Indiquer clairement si la mesure du critère de jugement est objective.

En cas d'utilisation d'un instrument de mesure, préciser s'il a été validé.

Inclure les définitions pour les résultats, dont l'interprétation est susceptible de varier, par exemple : définition de « hypertension. »

Dans les annexes, insérer une copie de tous les questionnaires et autres instruments de mesure.

### 1.6.3 Calendrier d'évaluation des résultats

On doit absolument déterminer le moment où il faudra les mesurer. En général, la période de suivi doit être suffisamment longue pour permettre d'observer une grande variété d'effets (autres que ceux qui sont immédiatement observés dès le début de l'intervention). On doit également inclure les mesures de base, et décider si les effets d'apprentissage et de perte des acquis doivent être évalués. Le moment de chaque mesure de résultat doit être rendu explicite.

### 1.6.4 Mesures de base

Les mesures de base incluent les caractéristiques démographiques de la population de l'étude, ainsi que les mesures de base de la performance.

Les interventions visant à modifier le comportement professionnel n'ont pas un effet immédiat et durable. Au contraire, leur effet est progressif (effet d'apprentissage), mais il décline avec le temps (effet de perte des acquis). Pour saisir et quantifier ces effets, il est important de mesurer les résultats par séries au fil du temps. Néanmoins, une telle décision dépend de l'équilibre entre les avantages et les coûts, en termes de ressources et de complexité de l'étude.

#### **Protocole :**

Spécifier la durée de la période de suivi.

Indiquez le moment où chaque résultat doit être mesuré (en incluant le calendrier pour toute mesure répétée).

### 1.6.5 Aspects statistiques : taille de l'échantillon et analyse statistique

La prise en compte des aspects statistiques impose la collaboration d'un méthodologiste ou statisticien connaissant les études en grappe.

La taille de l'échantillon retenue pour une étude doit être suffisante pour permettre de détecter une différence importante au point de vue clinique, sur le critère de jugement principal avec une probabilité souhaitée. Le calcul de la taille d'échantillon signifie l'équilibrage de la précision requise dans l'estimation de cette différence, et la minimisation de la taille d'échantillon requise pour obtenir cette précision.

Le calcul dépend de la manière dont l'étude a été conçue, du type de données des résultats primaires (binaires c/continues) et du niveau de signification souhaité, etc. Pour les essais en grappes, il faut tenir compte de l'effet de regroupement. Il faut également voir large, afin de tenir compte des désistements.

Il y a un choix à faire entre les difficultés logistiques et les coûts associés au recrutement de nouvelles grappes, et les difficultés associées à l'augmentation du nombre de patients par grappe.

**Protocole :**

Calculer la taille de l'échantillon, en se fondant sur le critère de jugement principal.

Ce calcul doit normalement refléter le type de résultat (ex. : binaire, continu).

Estimer les résultats prévus dans les groupes-intervention et témoins.

Indiquer le niveau (alpha) de signification, souvent estimé à 5 %.

Indiquer le niveau (bêta) d'efficacité, souvent estimé à 80 %.

Estimer le Coefficient de corrélation intraclasse (CCI).

Estimer la taille moyenne d'une grappe.

Augmenter la taille de l'échantillon en prévision de l'effet de mise en grappes.

Augmenter la taille de l'échantillon en prévision des désistements potentiels.

Donner la taille résultante de l'échantillon par groupe de comparaison.

### 1.6.6 Programmation de l'analyse statistique des résultats

Les méthodes statistiques standards, utilisées pour analyser les études randomisées sur des patients, se fondent sur un principe de base selon lequel le résultat obtenu pour un patient est complètement distinct de celui obtenu pour un autre patient. On dit qu'ils sont « indépendants ». Toutefois, ce principe ne s'applique pas aux études en grappes, car les patients au sein d'une même grappe vont vraisemblablement répondre de manière similaire. Dans le chapitre protocole, la section traitant de l'analyse statistique doit fournir des détails précis sur la manière dont l'effet grappe sera prise en compte.

Il faut également faire bien attention pour éviter ce que l'on appelle « l'erreur d'unité d'analyse. » Cette erreur résulte de l'usage inapproprié des mesures du niveau « patients », comme unité d'analyse servant à formuler des conclusions à propos d'un niveau supérieur, comme le niveau « professionnels de santé » ou les résultats combinés de l'interaction des patients et des professionnels. Son effet habituel est un accroissement erroné de l'efficacité d'une expérience, visant à détecter les différences entre les groupes intervention et contrôle (37-39).

L'analyse statistique des séries chronologiques fait appel à des méthodes statistiques particulières ([www.epoc.uottawa.ca](http://www.epoc.uottawa.ca)).

**Protocole :**

Fournir les références pour toute technique statistique d'analyse.

Donner des détails au sujet de l'analyse proposée pour les résultats primaires et secondaires.

Spécifier toute analyse de sous-groupes.

Fournir des informations sur le traitement des problèmes prévus, tels que les données manquantes.

Surtout, faire participer un statisticien à l'élaboration de cette section.

## 1.7 Aspects éthiques

Les principes éthiques généraux, et les normes qui s'appliquent aux études contrôlées et randomisées individuelles s'appliquent également aux ECR en grappes. Il existe cependant quelques considérations particulières qui s'appliquent à ce type d'études.

Les individus prenant part à une ECR en grappes ont une capacité limitée ou nulle à consentir à prendre part à l'étude, à recevoir l'intervention ou à quitter l'étude.

Edwards *et al* font la distinction, en termes moraux, entre deux types d'ECR en grappes (40) :

Études « grappe-grappe », où l'unité de répartition et d'intervention est la grappe, par exemple : instituer un système d'audit clinique ou projeter des vidéos à caractère pédagogique dans la salle d'attente d'un médecin.

Études « individu-grappe », où l'unité de répartition est la grappe, mais où l'unité d'intervention est l'individu ; par exemple l'évaluation d'une stratégie expérimentale de vaccination.

Dans une étude « grappe-grappe », où l'intervention cible l'ensemble de la grappe, le consentement individuel pour prendre part à l'étude, et recevoir l'intervention n'est pas applicable.

Pour une étude « individu-grappe », où l'intervention vise le participant individuel, le consentement ou le refus de recevoir l'intervention revient à l'individu.

En France, le terme « recherche biomédicale » ne s'applique pas :

- aux recherches dans lesquelles tous les actes sont pratiqués et les produits utilisés de manière habituelle, sans aucune procédure supplémentaire ou inhabituelle de diagnostic ou de surveillance ;
- aux recherches visant à évaluer les soins courants, autres que celles portant sur les médicaments, lorsque tous les actes sont pratiqués et les produits utilisés de manière habituelle, mais que des modalités particulières de surveillance sont prévues par un protocole, etc. (Loi n° 2004-806 du 9 août 2004, relative à la politique de santé publique art. 88 II [Journal Officiel du 11 août 2004], Loi n° 2006-450 du 18 avril 2006 pour la recherche art. 31 I [Journal Officiel du 19 avril 2006]).

Il est en général considéré que les études d'intervention sortent de ce cadre, mais il est recommandé de demander l'avis d'un comité de protection des personnes.

Par contre, les dispositions concernant les recherches qui comportent des données sensibles en termes de confidentialité (identité complète des personnes ou numéro collecté de Sécurité sociale) s'appliquent pleinement. Le responsable du fichier doit assurer la déclaration unitaire de la recherche auprès du Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé, puis de la Commission nationale de l'informatique et des libertés (CNIL).

## 1.8 Enregistrement de l'étude

Le fait de ne pas faire état des résultats d'une étude contrôlée randomisée est considéré de plus en plus comme une faute scientifique et éthique. Comme pour tout essai, les études d'intervention doivent désormais faire l'objet d'un enregistrement, qui est en particulier exigé par les éditeurs des principales revues médicales scientifiques (41).

Cette disposition ne s'applique pas pour les études quasi expérimentales non randomisées.

Pour savoir pourquoi et comment déclarer une étude randomisée contrôlée, il est préférable de se renseigner auprès d'une Direction régionale de la recherche clinique ou en consultant l'un des sites Internet suivants : [www.icmje.org](http://www.icmje.org), [www.clinicaltrials.gov](http://www.clinicaltrials.gov).

## 2. Durant la réalisation de l'étude

### 2.1 Maintien de l'intérêt des participants après le recrutement

Après avoir recruté des sujets pour l'étude, il est important d'élaborer des stratégies permettant d'assurer un suivi adéquat. Il s'agit là d'un point essentiel, notamment pour les études à période de recrutement longue. Dans un essai en grappes, un recrutement insuffisant dans un centre augmente de façon importante le risque de non comparabilité des groupes soumis à l'étude.

Quelle que soit la finalité de l'étude, la collecte des données n'est possible que si le personnel demeure activement impliqué pendant toute la durée de l'étude, et si les sujets manifestent leur volonté de poursuivre leur participation à l'étude.

Afin de s'assurer de leur engagement continu dans l'étude, il faut développer des stratégies de rétention postrecrutement (de préférence, dès la phase de préparation du protocole), en établissant une liste de toutes les activités devant être menées durant l'étude, afin d'atteindre ces objectifs. Il est important de les tenir informés des objectifs et/ou de l'état d'avancement de l'étude (réunions, entretiens personnels, documents d'information, site Internet dédié à l'étude, etc.)

Ceci exige que des ressources humaines et financières soient budgétisées.

#### **Protocole :**

Décrire la stratégie/procédure visant à maintenir l'intérêt du personnel de l'étude.

Décrire la stratégie/procédure visant à maintenir l'intérêt des participants de l'étude.

Expliquer les procédures pour tenir à jour les coordonnées des participants.

Expliquer les procédures de contrôle relatives au respect du protocole, aux écarts par rapport au protocole, à la divulgation des informations et aux cointerventions, pendant le déroulement de l'étude.

Expliquer les procédures de gestion des éventuelles pertes de grappes ou des abandons éventuels pendant l'étude, et identifier les mesures correctives si les objectifs de l'étude en dépendent.

### 2.2 Sécurité et effets indésirables

Il est important de garantir la sécurité et l'intégrité des patients participant à l'étude. Dans un essai contrôlé randomisé, l'investigateur principal est responsable de l'évaluation, de l'enregistrement, et de la notification des informations concernant tous les événements indésirables, liés au traitement ou à la procédure, ainsi que tout événement indésirable sérieux ou imprévu pouvant survenir en cours d'étude. Toutefois, étant donné la nature des études d'intervention, ce problème est moins préoccupant que pour les études portant sur les médicaments. La nature particulière de l'intervention doit permettre de décider si on doit ou non mettre en place un comité de contrôle de la sécurité et un système de notification des événements indésirables. En cas de doute, un avis peut être demandé à un méthodologiste spécialisé dans la réalisation d'essais contrôlés randomisés, à une unité de recherche clinique ou à une Direction régionale de recherche clinique.

#### **Protocole :**

Décrire les procédures de supervision des données et de la sécurité.

Mettre en place un comité indépendant (de préférence) de surveillance de la sécurité.

Justifier votre décision si vous décidez de ne pas avoir recours à un comité.

Établir (éventuellement) des plans d'analyse intermédiaires.

Envisager des règles statistiques d'arrêt pour l'efficacité et la sécurité.

## 2.3 Gestion des données

L'enregistrement de données constitue probablement l'étape la plus cruciale de tout le processus de gestion des données. Les erreurs commises à ce niveau peuvent s'avérer plus difficiles à détecter et à corriger que celles commises au cours d'une étape ultérieure. Les procédures de gestion des données utilisées au cours d'étapes ultérieures n'auront que peu d'effet sur la fiabilité et la validité des données source.

### 2.3.1 Collecte des données

La première étape, dans l'élaboration de la stratégie de collecte des données, consiste à choisir les données à collecter et le moment où cette collecte doit être. Toutes les décisions doivent être motivées par les objectifs scientifiques de la recherche, et il faut veiller à ne collecter que les données nécessaires pour atteindre les objectifs.

Le recueil de données peut être effectué au moyen d'un questionnaire, de systèmes électroniques d'enregistrement de données ou de données collectées de façon systématique (cahiers de recueil spécifiques).

#### **Protocole :**

Décrire les données qui vont être collectées, et indiquer les moments auxquels ces données seront collectées.

Décrire brièvement la stratégie de collecte des données.

Indiquer la manière dont seront traités : l'administration du questionnaire, les besoins en personnel et en formation, la transmission des données au centre de coordination, la sécurité et la confidentialité.

Inclure en annexe un calendrier des visites aux patients ou aux médecins ou aux centres, et une description de ce qui sera fait à chaque visite.

Inclure en annexe tous les formulaires servant à collecter les données.

Identifier celui ou ceux qui se chargeront des questions relatives à la protection des données.

### 2.3.2 Gestion des données

Les études doivent, autant que possible, utiliser un système informatisé pour la gestion des données. Les systèmes informatiques sont utiles pour suivre l'évolution du flux de participants et de grappes durant les études (par exemple pour le recrutement, les désistements, les raisons des désistements, la collecte de données et l'enregistrement du flux d'informations entre vous et les participants ou les grappes, comme l'envoi et le retour de questionnaires). Ces systèmes permettent aussi une validation rapide et efficace des données, ainsi qu'un contrôle de la qualité.

#### **Protocole :**

Décrire les procédures qui seront incorporées dans le système de gestion des données (par exemple, enregistrement de l'envoi/retour de questionnaires).

Indiquer les noms des logiciels de saisie et de gestion, ainsi que des langages de programmation qui seront utilisés pour l'étude.

Détailler la configuration informatique minimale, nécessaire pour la gestion et la transmission de données.

### 2.3.3 Saisie des données

La saisie des données est le processus de conversion des informations, quelle qu'en soit la source originale, sous forme de données sur support électronique pouvant être gérées et



analysées au moyen d'un ordinateur. La méthode la plus courante de saisie de données est la transcription d'informations contenues sur un formulaire papier, sur un fichier informatique. La saisie des données peut être effectuée par une seule personne (saisie simple) ou deux fois par des personnes différentes (saisie double). Le logiciel servant à stocker et gérer les données est appelé système de gestion de base de données.

Récemment, la technologie moderne a permis de remplacer la phase de saisie manuelle par des systèmes de « capture de données ». Grâce à cette méthode, les données sont soit scannées par voie électronique, soit saisies directement dans un ordinateur, par exemple au moyen d'un ordinateur de poche (PDA) ou *via* Internet.

**Protocole :**

Décrire succinctement la manière dont les données seront saisies dans le système de gestion des données.

Décider si cette tâche sera effectuée uniquement par le centre de coordination des données, et préciser si les données seront saisies manuellement (saisie simple ou double) ou bien scannées.

### 2.3.4 Contrôle de qualité

Le succès de toute étude clinique dépend de la qualité des données collectées. Pour garantir l'obtention de données de bonne qualité, il est nécessaire d'élaborer des procédures d'assurance qualité complètes et détaillées, telles que celles élaborées pour le contrôle des centres cliniques et celles servant à assurer la qualité du traitement des données au centre de coordination. De même que pour tous les aspects des études cliniques, un programme de contrôle-qualité nécessite un souci du détail et une planification minutieuse. Ce programme doit inclure un plan de vérification des données (Quelle quantité sera vérifiée et par qui ? selon quelle méthode ?), des listes de contrôle (registre détaillé montrant quelles données ont été modifiées, la raison de la modification, qui l'a effectuée et quand ?), et des procédures de vérification des sources (Quelle quantité de données seront comparées au document source ?).

**Protocole :**

Décrire les mesures permettant de garantir que la manipulation ou la transmission des données n'entraîne pas de modifications involontaires.

Ces mesures doivent inclure : la vérification des données, les listes de contrôle et les procédures de vérification des données source.

## 3. Après la réalisation de l'étude

### 3.1 Rapport final

À la fin d'une période de subvention ou de financement, on est souvent amené à rédiger et à soumettre un rapport final et un bilan financier. La date de présentation est souvent située à quelques mois de la date d'échéance du financement. Les informations requises peuvent inclure :

- des réalisations scientifiques, telles que des publications ;
- la méthodologie employée ;
- les progrès scientifiques réalisés (ainsi que toute modification qui en découle au niveau des objectifs de l'étude) ;
- l'impact réel et potentiel sur la santé et la qualité de vie ;
- les efforts entrepris pour la diffusion des résultats.

## 3.2 Politique de publication

Dans le protocole de l'étude, il faut exposer clairement et documenter la politique de publication à la fois pour les rapports intermédiaires et les articles scientifiques finaux. Aucune publication, qu'elle soit écrite ou orale, ne doit être faite avant que le manuscrit définitif ait été approuvé et sa publication acceptée. Il faut également obtenir l'approbation préalable de toutes les parties prenantes ayant participé à l'étude. Il est courant de mettre en place un « comité de rédaction. »

Les dispositions concernant la paternité de l'œuvre doivent avoir été approuvées et documentées dans le protocole. Beaucoup d'études de grande envergure ont recours à la « paternité de groupe », et listent à la fin de leur publication les détails de la contribution de chacun à chaque stade de l'étude, sans oublier le personnel clinique ayant collaboré.

Un format de rapport courant pour les essais contrôlés en grappes est désormais disponible. Il est connu sous le nom de *CONSORT statement : extension to cluster randomised trials* (énoncé CONSORT : extension aux études randomisées en grappes), et fournit une liste de contrôle contenant 22 points portant principalement sur les méthodes, les résultats et les discussions d'un rapport sur une étude contrôlée et randomisée en grappes. Il inclut les informations clés nécessaires à l'évaluation de la validité interne et externe de l'étude (42-44).

### **Protocole :**

Décrire les règles et recommandations régissant la publication des résultats techniques et scientifiques de l'étude.

S'assurer que ces règles visent à promouvoir l'exactitude scientifique et technique de toute publication, et qu'elles garantissent que tous les enquêteurs et autres personnes ayant contribué à l'étude seront mentionnés.

## Références

1. Grimshaw G, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, *et al.* Effectiveness and efficiency of guidelines dissemination and implementation strategies. *Health Technol Assess* 2004;8(6).
2. Agence nationale d'accréditation et d'évaluation en santé. Efficacité des méthodes de mise en œuvre des recommandations médicales. Paris: ANAES; 2000.
3. Durand-Zaleski I, Colin C, Blum-Boisgard C. An attempt to save money by using mandatory practice guidelines in France. *BMJ* 1997;315(7113):943-6.
4. Durieux P, Chaix-Couturier C, Durand-Zaleski I, Ravaud P. From clinical recommendations to mandatory practice. The introduction of regulatory practice guidelines in the French healthcare system. *Int J Technol Assess Health Care* 2000;16(4):969-75.
5. Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003;12(1):47-52.
6. Kosecoff J, Kanouse DE, Rogers WH, McCloskey L, Winslow CM, Brook RH. Effects of the National Institute of Health Consensus Development Program on physician practice. *JAMA* 1987;258(19):2708-13.
7. Kucher N, Koo S, Quiroz R, Cooper JM, Paterno MD, Soukonnikov B, *et al.* Electronic alerts to prevent venous thromboembolism among hospitalized patients. *N Engl J Med* 2005;352(10):969-77.
8. Diwan VK, Eriksson B, Sterky G, Tomson G. Randomization by group in studying the effect of drug information in primary care. *Int J Epidemiol* 1992;21(1):124-30.
9. Donner A, Birkett N, Buck C. Randomization by cluster. Sample size requirements and analysis. *Am J Epidemiol* 1981;114(6):906-14.
10. Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. *Clin Trials* 2005;2(2):99-107.
11. Eldridge SM, Ashby D, Feder GS, Rudnicka AR, Ukoumunne OC. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clin Trials* 2004;1(1):80-90.
12. Guittet L, Giraudeau B, Ravaud P. A priori postulated and real power in cluster randomized trials: mind the gap. *BMC Med Res Methodol* 2005;5:25.
13. Auleley GR, Ravaud P, Giraudeau B, Kerboull L, Nizard R, Massin P, *et al.* Implementation of the Ottawa ankle rules in France. A multicenter randomized controlled trial. *JAMA* 1997;277(24):1935-9.
14. Horbar JD, Carpenter JH, Buzas J, Soll RF, Suresh G, Bracken MB, *et al.* Collaborative quality improvement to promote evidence based surfactant for preterm infants: a cluster randomised trial. *BMJ* 2004;329(7473):1004.
15. Thomas RE, Croal BL, Ramsay C, Eccles M, Grimshaw J. Effect of enhanced feedback and brief educational reminder messages on laboratory test requesting in primary care: a cluster randomised trial. *Lancet* 2006;367(9527):1990-6.
16. Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, *et al.* Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325(7370):941.
17. Ricordeau P, Durieux P, Weill A, Chatellier G, Vallier N, Bissery A, *et al.* Effect of a nationwide program of educational outreach visits to improve the processes of care for patients with type 2 diabetes. *Int J Technol Assess Health Care* 2003;19(4):705-10.
18. Durieux P, Nizard R, Ravaud P, Mounier N, Lepage E. A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior. *JAMA* 2000;283(21):2816-21.
19. Ray-Coquard I, Philip T, de Laroche G, Froger X, Suchaud JP, Voloch A, *et al.* A controlled "before-after" study: impact of a clinical guidelines programme and regional cancer network organization on medical practice. *Br J Cancer* 2002;86(3):313-21.
20. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;283(13):1715-22.
21. Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, *et al.* Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med* 2004;141(10):771-80.
22. Luck J, Peabody JW, Lewis BL. An automated scoring algorithm for computerized clinical vignettes: evaluating physician performance against explicit quality criteria. *Int J Med Inform* 2006;75(10-11):701-7.
23. Veloski J, Tai S, Evans AS, Nash DB. Clinical vignette-based surveys: a tool for assessing physician practice variation. *Am J Med Qual* 2005;20(3):151-7.

24. Norcini J. Back to the future: clinical vignettes and the measurement of physician performance. *Ann Intern Med* 2004;141(10):813-4.
25. Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ* 1995;311(7008):793-6.
26. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;283(13):1715-22.
27. Durieux P, Ravaud P, Porcher R, Fulla Y, Manet C, Chaussade S. Long-term impact of a restrictive laboratory test ordering form on tumor marker prescriptions. *Int J Technol Assess Health Care* 2003;19(1):106-13.
28. Varroud-Vial M, Simon D, Attali J, Durand-Zaleski I, Bera L, Attali C, *et al.* Improving glycaemic control of patients with Type 2 diabetes in a primary care setting: a french application of the staged diabetes management programme. *Diabet Med* 2004;21(6):592-8.
29. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. *MJA* 2004;(180):S57-S60.
30. Saillour-Glenisson F, Michel P. Facteurs individuels et collectifs associés à l'application des recommandations de pratique clinique par le corps médical. *Revue de littérature. Rev Epidemiol Santé Publ* 2003;51:65-80.
31. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, *et al.* Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;282(15):1458-65.
32. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362(9391):1225-30.
33. Grimshaw J, McAuley LM, Bero LA, Grilli R, Oxman AD, Ramsay C, *et al.* Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Qual Saf Health Care* 2003;12(4):298-303.
34. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. *Qual Saf Health Care* 2003;12(1):40-6.
35. Agence nationale d'accréditation et d'évaluation en santé. Construction et utilisation des indicateurs dans le domaine de la santé. Principes Généraux. Paris: ANAES; 2002.
36. Coordination pour la Mesure de la Performance et l'Amélioration de la Qualité Hospitalière. Pratique clinique: prise en charge hospitalière de l'infarctus du myocarde après la phase aiguë. Synthèse des résultats. Le Kremlin Bicêtre: Projet COMPAQH; 2006.
37. Whiting-O'Keefe QE, Henke C, Simborg DW. Choosing the correct unit of analysis in Medical Care experiments. *Med Care* 1984;22(12):1101-14.
38. Campbell MK, Mollison J, Steen N, Grimshaw JM, Eccles M. Analysis of cluster randomized trials in primary care: a practical approach. *Fam Pract* 2000;17(2):192-6.
39. Kerry SM, Bland JM. Analysis of a trial randomised in clusters. *BMJ* 1998;316(7124):54.
40. Edwards SJ, Braunholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the design and conduct of cluster randomised controlled trials. *BMJ* 1999;318(7195):1407-9.
41. Krleza-Jeric K, Chan AW, Dickersin K, Sim I, Grimshaw J, Gluud C. Principles for international registration of protocol information and results from human trials of health related interventions: Ottawa statement (part 1). *BMJ* 2005;330(7497):956-8.
42. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med Res Methodol* 2001;1:2.
43. Campbell MK, Elbourne DR, Altman DG, Consort Group. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;328(7441):702-8.
44. Uniform requirements for manuscripts submitted to biomedical journals. International Committee of Medical Journal Editors. *Ann Intern Med* 1997;126(1):36-47.

# ANNEXE 1 : *Cochrane Effective Practice and Organisation of Care Group (EPOC)*

Types d'intervention (source = [www.epoc.uottawa.ca](http://www.epoc.uottawa.ca))

## 1. Interventions orientées vers les professionnels de santé

- *Distribution de supports didactiques* : distribution de recommandations de pratique publiées ou imprimées, de supports audiovisuels ou de publications sur support électronique. Les documents peuvent être fournis à titre personnel ou par mailing de masse.
- *Conférences* : participation des professionnels de santé à des conférences, à des exposés, à des ateliers ou à des sessions de formation en dehors de leurs propres activités professionnelles.
  - Conférences en groupes restreints (participation active) ;
  - Conférences en groupes élargis (participation passive).
- *Processus consensuels locaux* : implication des professionnels afin qu'ils conviennent de l'importance du thème clinique choisi et de la pertinence de la méthode utilisée pour le gérer.
- *Visites à domicile* : désignation d'une personne chargée de rencontrer les professionnels de santé dans leurs locaux professionnels dans le but de leur fournir des informations. Les informations fournies peuvent inclure des retours d'informations sur les performances des professionnels.
- *Leaders d'opinion locaux* : intervention faisant appel à des prestataires reconnus par leurs collègues pour « leur influence didactique. » Il est souhaitable que la manière dont les leaders d'opinion ont été identifiés (par leurs collègues) et recrutés soit clairement indiquée.
- *Interventions médiatisées par les patients* : toute intervention visant à faire évoluer les performances des professionnels de santé, lorsque les informations ont été recherchées par les patients ou qu'elles leur ont été directement fournies ; par exemple, envois directs de courriers aux patients, conseils donnés aux patients par des tiers, documents d'information fournis aux patients ou déposés dans les salles d'attente.
- *Audit et rétroaction (« feedback »)* : toute information ou tout résumé sur des performances cliniques en soins de santé pendant une période spécifiée. Les informations peuvent être fournies oralement ou par écrit ; elles peuvent également inclure des recommandations pour une action clinique. Les informations sur les performances des prestataires peuvent provenir de dossiers médicaux, de bases de données informatiques, de faits observés ou de patients.
  - Audit interne : un audit réalisé par les professionnels de santé eux-mêmes.
  - Audit externe : les professionnels de santé obtiennent des informations sur leurs propres performances à partir de sources externes. Les interventions suivantes **ne** devraient être incluses **ni** dans l'audit **ni** dans la rétroaction :
    - Diffusion d'informations cliniques qui ne reflètent pas directement les performances des prestataires, et qui sont directement recueillies par les enquêteurs auprès des patients, ex. : notation d'un instrument d'évaluation de l'état dépressif.
    - *feedback* obtenu à partir d'informations issues de dossiers de santé de patients individuels, dans un format différent (ex. : informatique).
- *Rappels* : Toute intervention, manuelle ou automatisée, qui incite le prestataire de soins de santé à exécuter une action clinique.

Cela comprend les interventions suivantes :

  - Aide automatisée à la décision. (Utilisation d'un système de connaissance actif, qui se sert d'au moins deux éléments d'information relatifs à un patient pour générer un avis médical spécifique).
  - Rapports concomitants. (Destinés aux professionnels de santé lors d'une rencontre de sensibilisation, pour leur rappeler les actions nécessaires à entreprendre pour des patients particuliers).
  - Rappels entre visites. (Destinés aux prestataires entre les visites lorsqu'il y existe une preuve démontrant que les soins fournis à des patients spécifiques sont sous-optimaux,

comme par exemple, lorsque le résultat d'un test est anormal et que l'on ne retrouve pas le suivi approprié dans le dossier médical).

- Rapport de laboratoire amélioré (Rapport de laboratoire consécutif à un résultat anormal et destiné à des prestataires. Il contient des informations supplémentaires sur des recommandations de suivi spécifique).
- Soutien administratif (système de rendez-vous de suivi ou autocollants sur dossiers).
- Rappels implicites (valeurs prédictives pour des résultats de tests anormaux sans qu'il y ait une recommandation explicite pour une action).
- *Interventions personnalisées* : Utilisation d'entretiens personnels, de discussions de groupe (« groupes de réflexion ») ou d'enquêtes sur les prestataires, pour identifier les obstacles au changement et l'élaboration ultérieure d'une intervention visant à lever les obstacles identifiés.
- Revue par les pairs.
- *Stratégies combinées* : spécifier l'utilisation du classement des interventions ci-dessus, le cas échéant.
- *Autre*, spécifier.

## 2. Interventions financières

- Interventions sur les prestataires :
  - Rémunération à l'acte.
  - Capitation (le prestataire perçoit un montant fixe par personne dans la population-cible par unité de temps pour fournir des soins spécifiques).
  - Salariat (un prestataire perçoit un salaire pour fournir des soins de santé spécifiques).
  - Rémunération au rendement d'un prestataire (un prestataire particulier reçoit une rétribution ou un avantage financier direct ou indirect pour l'exécution d'une action spécifique).
  - Rémunération au rendement d'une institution (une institution ou des groupes de prestataires reçoivent des rétributions ou des avantages financiers directs ou indirects pour l'exécution d'une action spécifique).
  - Subvention/indemnité versée à un prestataire (un prestataire particulier reçoit une rétribution ou un avantage financier direct ou indirect, non lié à l'exécution d'une action spécifique).
  - Subvention/indemnité versée à une institution (une institution ou un groupe de prestataires reçoivent une rétribution ou un avantage financier direct ou indirect, non lié à l'exécution d'une action spécifique).
  - Pénalité imposée à un prestataire (une pénalité financière directe ou indirecte est imposée à un prestataire particulier pour sanctionner une action inappropriée).
  - Pénalité imposée à une institution (une pénalité financière directe ou indirecte est imposée à une institution ou à des groupes de prestataires pour sanctionner une action inappropriée).
  - Changements dans le formulaire (ajouts ou retraits de produits disponibles remboursables).
  - Autre, spécifier.
- Interventions sur les patients :
  - Prime.
  - Copaiement.
  - Ticket modérateur.
  - Rémunération des patients.
  - Subvention/indemnité versée aux patients.
  - Pénalité imposée aux patients.
  - Autre, spécifier.

## 3. Interventions organisationnelles

Ces interventions peuvent inclure des changements au niveau de l'aménagement matériel des centres de soins de santé, des systèmes de gestion des dossiers médicaux ou de la propriété.

- Interventions structurelles :
  - Changements au niveau des locaux/sites de prestation de soins (ex. : déménagement d'un service d'orthogénie précédemment installé dans un hôpital vers un établissement scolaire).
  - Télémédecine (fourniture de moyens de communication et discussion de cas entre des professionnels de santé éloignés).
  - Changements au niveau des systèmes de gestion des dossiers médicaux (ex. : passage des dossiers sur support papier aux dossiers sur support informatique).
  - Autres changements au niveau des systèmes de mise à jour ou de recherche d'informations (ex. : système de suivi de patients).
  - Autres changements au niveau de l'aménagement matériel, des installations et des équipements.
  - Changements au niveau de l'étendue et/ou de la nature des services (ex. : mise en place d'un service de chirurgie ambulatoire).
  - Changements au niveau de la présence et de l'organisation des processus de gestion de la qualité.
  - Changements au niveau de la propriété et/ou de l'affiliation des hôpitaux et autres équipements.
  - Autres changements structurels au niveau de l'organisation, personnel non compris (spécifier).
- Interventions axées sur le personnel de santé :
  - Examen des tâches des professionnels de santé (Changements au niveau de l'étendue des tâches parmi les professionnels de santé également appelé « substitution professionnelle » ou « empiètement sur les limites » ; ex. : des infirmières sages-femmes assurant des soins obstétricaux ou des pharmaciens donnant des conseils sur les médicaments, alors que ces fonctions étaient auparavant assurées par des infirmières et des médecins).
  - Équipes multidisciplinaires (des professionnels de santé appartenant à différentes disciplines travaillent ensemble en équipe pour fournir des soins de santé à un patient ou à une population).
  - Gestion de cas (un professionnel de santé se charge de coordonner les soins fournis à un patient par plusieurs prestataires et/ou centres de soins).
  - Autre intégration de services (mécanismes de suivi pour coordonner les soins fournis à un patient par « franchissement des frontières organisationnelles ou hospitalières », parfois appelés « soins intégrés »).
  - Interventions au niveau de la compétence collective (modifications portant sur le nombre, le type ou la qualification des membres du personnel).
  - Interventions visant à améliorer les conditions de travail et la gratification matérielle/morale des prestataires (ex. : interventions pour « remonter le moral »).
  - Autre, spécifier.
- Interventions axées sur les patients :
  - Interventions facilitant la participation de patients individuels (ex. : outils d'aide à la décision destinés aux patients).
  - Interventions facilitant la participation de groupes de patients (ex. : groupes de réflexion, panels de patients).
  - Autre, spécifier.

#### **4. Interventions réglementaires**

Toute intervention visant à modifier la délivrance ou les coûts des services de santé par le biais d'un règlement ou d'une loi. Il est possible qu'il y ait un chevauchement entre ces interventions et des interventions organisationnelles et financières, ou qu'une intervention comporte des éléments provenant de plusieurs catégories.

- Évolutions de la responsabilité médicale.
- Gestion des réclamations des patients.
- Accréditation.
- Autorisation d'exercer.
- Autre, spécifier.

## ANNEXE 2 : Grille de choix des thèmes de travail

### **1. Objectif**

- Améliorer la qualité des soins délivrés aux patients (+ /- impact sur les coûts) 2
- Réduire les coûts 1
- Pas d'objectif précis (par ex. : intérêt de l'évaluateur) 0

### **2. Caractère prioritaire**

- Obligation réglementaire ou le thème s'inscrit dans le cadre de l'accréditation 2
- Priorités locale, régionale ou nationale 1
- Pas de caractère prioritaire explicite 0

### **3. Transversalité**

- Thème transversal multiprofessionnel 2
- Thème transversal monoprofessionnel 1
- Le thème n'est pas transversal 0

### **4. Référentiel**

- Un référentiel valide existe 2
- Un référentiel sera simple à élaborer à partir de recommandations existantes 1
- Des recommandations devront être élaborées préalablement 0

### **5. Mise en œuvre**

- Une stratégie de mise en œuvre simple est possible 2
- Une stratégie de mise en œuvre est possible, mais difficile à réaliser 1
- Pas de stratégie de mise en œuvre possible 0

### **6. Évaluation**

- L'information pour évaluer la mise en œuvre du thème existe (indicateur, etc.) 2
- L'information n'existe pas, mais est possible à obtenir 1
- Information difficile à mettre en œuvre 0

### **7. Appréciation globale**

- L'impact prévisible est important, le projet est pérenne, l'investissement est acceptable 2
- Un des trois éléments précédents existe 1
- Aucun des trois éléments précédents n'existe 0

La cotation de cette grille est donnée à titre indicatif. Mais une cotation 0 à l'une des 7 questions exclut le thème évalué.



## ANNEXE 3 : Analyse critique d'une étude d'intervention

(Data collection check list, source [www.epoc.uottawa.fr](http://www.epoc.uottawa.fr)).

(Essai contrôlé, étude avant/après avec groupe-témoin ou série chronologique avec un critère de jugement objectif de mesure de performance, de comportement ou de résultat).

### DESCRIPTION GÉNÉRALE DE L'ÉTUDE

Référence de l'article

**Titre**

**Revue**

**Année de publication**

**Auteurs**

Objectifs

**Quelle est la nature du changement souhaité ?**

Amélioration des pratiques	Oui	Non
Introduction d'une nouvelle technologie	Oui	Non
Contrôle des coûts	Oui	Non
Autre	Oui	Non

Plan de l'étude

**Quel est le schéma de l'étude ?**

Essai contrôlé randomisé	Oui	Non
Simple		
Complexe		
Blocs incomplets		
Carrés latins		
Plans factoriels		
Étude avant/après avec groupe-témoins	Oui	Non
Série chronologique	Oui	Non

**Quelle est la nature du groupe-témoin ?**

Pas d'intervention  
Pratique habituelle  
Autre intervention

Intervention

**Quel est le type d'intervention soumise à l'étude ?**

Intervention professionnelle	Oui	Non
FMC (formation médicale continue)	Oui	Non
Audit et retour d'information	Oui	Non

Diffusion simple de documents d'information	Oui	Non
Leaders d'opinion	Oui	Non
Rappels au moment de la décision	Oui	Non
Visites à domicile	Oui	Non
Action sur les patients	Oui	Non
Mass media	Oui	Non
Consensus locaux	Oui	Non
Marketing	Oui	Non
Autre	Oui	Non
Intervention organisationnelle	Oui	Non
Orientée vers le professionnel		
Orientée vers le patient		
Modifications structurelles		
Intervention financière		
Orientée vers le professionnel		
Orientée vers le patient		
Orientée vers l'institution		

### **Support de l'intervention**

Informatique	Oui	Non
Papier	Oui	Non
Oral	Oui	Non
Oral direct	Oui	Non
Oral téléphonique	Oui	Non
Autre	Oui	Non

### **Quelle est la nature de la référence de pratique éventuellement utilisée ?**

Recommandation de pratique clinique  
Conférence de consensus  
Consensus formalisé d'experts  
Standard de pratique  
Règle de décision  
Aucune

### **Cette référence de pratique est-elle basée sur la preuve scientifique ?**

Oui Non Pas clair

### Lieu/Contexte

### **Quel est le type de pratique concernée ?**

Prévention

Diagnostic  
Demande d'examen  
Prescription  
Prise en charge globale d'un problème  
Dossier médical  
Organisation des soins  
Résultat des soins  
Dépistage  
Pronostic  
Autre

**Quelles sont les informations sur le contexte dans lequel l'étude a été réalisée ?**

Lieu

Pays

Région

Ville

Type de soins

Ambulatoires

Hospitaliers

Statut académique

Public

Oui Non

Privé

Oui Non

Universitaire

Oui Non

Inconnu

Oui Non

Mode de prise en charge (remboursement) des soins

Connu

Inconnu

**À quelle date l'étude a-t-elle été réalisée ? (date du début de l'intervention)**

Population

**Quelles sont les informations sur les professionnels concernés par l'intervention ?**

Profession

Médecin

Oui Non

Généraliste

Oui Non

Spécialiste

Oui Non

En formation

Oui Non

Infirmière

Oui Non

Autre

Oui Non

Nombre

**Quelles sont les caractéristiques des patients concernés ?**

Problème clinique

Âge moyen

Nombre total, lors de l'intervention

Critères de jugement

**Combien y a-t-il de critères de jugement ? (n < 4, par ex.)**

**Pour chaque critère de jugement :**

**Que concerne ce critère de jugement ?**

La procédure

Mesure du comportement du médecin

Mesure du comportement du patient

Les résultats

Morbidité des patients

Mortalité des patients

Autre

**Quelle est l'unité d'analyse correspondante ?**

Patient

Professionnel

Cabinet

Hôpital

Institution

Population

Autre

Aspects éthiques

Un **comité d'éthique** a-t-il approuvé l'étude ?

Oui Non

**CRITÈRES DE QUALITÉ**

**Connaît-on le pourcentage de professionnels de santé (ou unités d'allocation) concernés par l'étude rapporté au nombre total de professionnels de la population-cible ?**

Oui Non Pas clair

**Quelle est l'unité d'allocation des groupes ? (unité de randomisation)**

Patient

Professionnel

Cabinet

Hôpital

Institution

Population

Autre  
Aucune

### Quelles sont les unités d'analyse ?

Patient	Oui	Non
Professionnel	Oui	Non
Cabinet	Oui	Non
Hôpital	Oui	Non
Institution	Oui	Non
Population	Oui	Non
Autre	Oui	Non

### Critères de qualité du schéma expérimental

#### **Si essai contrôlé randomisé :**

Dissimulation de la randomisation (ECR)                      Oui    Non    Pas clair  
(insu/l'aveugle/allocation concealment)

*Notez OUI si*

*L'unité de randomisation est l'institution, le groupe de professionnels, le professionnel et la procédure de randomisation est explicitement décrite, par exemple table de nombres aux hasard ;*

*l'unité de randomisation est le patient ou l'épisode de soins, et il existe une forme de randomisation centralisée ou sur site informatisée ou par enveloppes scellées opaques.*

*Notez PAS CLAIR si*

*L'unité de randomisation n'est pas décrite explicitement ;*

*l'unité de randomisation est le patient ou l'épisode de soins et les auteurs décrivent avoir utilisé une liste, une table de randomisation, des enveloppes ou enveloppes scellées.*

*Notez NON si*

*Les auteurs déclarent avoir utilisé une forme d'allocation alternée comme le numéro de dossier, la date de naissance, le jour d'inclusion (« controlled clinical trial » dans la classification EPOC) ;*

*l'unité de randomisation est le patient ou l'épisode de soins et le processus d'allocation est totalement transparent ;*

*Le processus d'allocation de sujets a été modifié, par les investigateurs, les professionnels de santé, les patients.*

#### **Si étude avant/après avec groupe-témoin :**

Le recueil de données dans les groupes d'étude est contemporain	Oui	Non	Pas clair
Les groupes sont comparables	Oui	Non	Pas clair

#### **Si série chronologique :**

Le début de l'intervention est clairement défini	Oui	Non	Pas clair
--	-----	-----	-----------

Protection contre un changement séculaire

(non lié à l'intervention)

Oui Non Pas clair

L'intervention est indépendante d'autres changements

Il y a suffisamment de points de mesure pour une analyse statistique

Il y a un test statistique de tendance (analyse de rupture de pente)

*Notez OUI si*

*Il existe au moins 20 points de mesure avant intervention, **et** les auteurs ont réalisé une analyse classique type modèle ARIMA*

*Il existe au moins 3 points de mesure avant et après intervention, **et** les auteurs ont réalisé une analyse de mesures répétées*

*Il existe au moins 3 points de mesure avant et après intervention, **et** les auteurs ont réalisé une analyse ANOVA « or multiple t-tests », **et** il existe au moins 30 observations par point*

*Notez PAS CLAIR si*

*Le nombre de points n'est pas clairement spécifié dans le texte ou sur les tableaux ou figures (sera noté NON si cette information ne peut être obtenue des auteurs*

*Notez NON si*

*Aucune des conditions précédentes n'est rencontrée*

*l'intervention n'a pas affecté le recueil de données*

Oui Non Pas clair

**Pour toutes les études :**

> 80 % des professionnels suivis

Oui Non Pas clair

> 80 % des patients suivis

Oui Non Pas clair

Critère(s) de jugement évalué(s) en aveugle ou

critère(s) de jugement objectif(s)

Oui Non Pas clair

Critère(s) de jugement sûr (90 % d'accord

entre deux lecteurs ou kappa > 0,8) ou

critère(s) de jugement objectif(s)

Oui Non Pas clair

Mesure des pratiques avant intervention et

il n'existe pas de différence entre groupes

Oui Non Pas clair

Protection contre une contamination entre groupes

Oui Non Pas clair

*Notez OUI si allocation par institution, groupe de médecins, et s'il est très peu probable que le groupe-témoin ait reçu l'intervention*

*Notez PAS CLAIR si les professionnels ont été randomisés à l'intérieur d'une structure ou d'un groupe de professionnels, et qu'il est possible qu'il y ait eu communication entre groupe expérimentale et groupe-témoins.*

*Notez NON s'il est probable que le groupe-témoin ait reçu l'intervention (par exemple, si étude en « cross over » ou si randomisation par patients)*

## RÉSULTATS

### Quels sont les principaux résultats ?

Quels sont les principaux résultats ?

Indicateur de résultat dans chaque groupe (témoin et intervention), +/- IC (RR)

Indicateur de comparaison relative, +/- IC (RRR)

Indicateur de comparaison absolue, +/- IC (RAR)

A-t-on comparé la différence pré/postintervention dans les différents groupes soumis à l'étude ?  
Oui Non

Quelle interprétation en faites-vous (importance de l'effet) ?

Valeur de la significativité statistique p

Interprétation : significativité clinique

**Quelle est la durée du suivi après mise en place de l'intervention ? (début)**

**Quelle est la durée du suivi après intervention ? (fin)**

**A-t-on identifié un possible « effet plafond » ? (effet max = > plateau)**

Oui Non

**Remarque** : ensemble de réponses o/n = non exclusive

## **Auteurs**

D<sup>r</sup> Pierre Durieux, Service de Santé Publique, Hôpital européen Georges-Pompidou, Paris.

P<sup>r</sup> Philippe Ravaud, INSERM U738, Dpt d'épidémiologie, de biostatistique et de recherche clinique, Groupe hospitalier Bichat-Claude-Bernard, Paris.

## **Groupe de lecture**

P<sup>r</sup> Serge Briançon, Santé Publique, Nancy.

D<sup>r</sup> Bernard Burnand, Médecine sociale et préventive, Lausanne.

P<sup>r</sup> Isabelle Durand-Zaleski, Santé Publique, Créteil.

D<sup>r</sup> Hector Falcoff, Médecin généraliste, Paris.

P<sup>r</sup> Patrice François, Santé Publique, Grenoble.

D<sup>r</sup> Dominique Huas, Médecin généraliste, Nanterre.

M<sup>me</sup> Annick Macrez, Cadre infirmier, Paris.

D<sup>r</sup> Alexandre Pariente, Gastro-entérologue, Pau.

P<sup>r</sup> Isabelle Ray-Coquard, Oncologue, Lyon.

P<sup>r</sup> Pierre-Marie Roy, Anesthésie-réanimation, Angers.

D<sup>r</sup> Florence Saillour-Glenisson, Santé publique, Pessac.

M<sup>me</sup> Sophie Saisy-Callaert, Pharmacienne, Créteil.

D<sup>r</sup> Louis-Rachid Salmi, Santé Publique, Bordeaux.



## Fiche descriptive

<b>TITRE</b>	<b>Évaluer l'impact d'interventions destinées à améliorer les pratiques cliniques</b>
<b>Méthode de travail</b>	Guide Méthodologique
<b>Date de mise en ligne</b>	Octobre 2007
<b>Objectifs</b>	Présenter les différentes caractéristiques méthodologiques d'une étude d'intervention destinée à améliorer la qualité des soins, et les différentes étapes de sa réalisation
<b>Professionnels concernés</b>	Sociétés savantes, organismes et groupes de professionnels souhaitant évaluer l'impact d'une intervention en santé
<b>Demandeur</b>	Haute Autorité de santé
<b>Promoteur</b>	Haute Autorité de santé Coordination : D <sup>r</sup> N. Mlika-Cabanne, Adjoint au chef du service des recommandations professionnelles de la HAS (Directeur de l'évaluation des stratégies de santé de la HAS : D <sup>r</sup> P. Michel)
<b>Pilotage du projet</b>	Secrétariat : M <sup>mes</sup> Sladana Praizovic et Marilyn Rubinsztein Recherche documentaire : M <sup>me</sup> Emmanuelle Blondet, avec l'aide de M <sup>me</sup> Maud Lefevre, service de documentation de la HAS (Chef de service : M <sup>me</sup> Frédérique Pajès)
<b>Participants</b>	Auteurs Groupe de lecture
<b>Recherche documentaire</b>	De janvier 2000 à juin 2007
<b>Auteurs du document</b>	D <sup>r</sup> Pierre Durieux P <sup>r</sup> Philippe Ravaud
<b>Validation</b>	Avis de la Commission Évaluation des stratégies de santé de la HAS. Validation par le Collège de la HAS en juin 2007.
<b>Format</b>	Téléchargeables gratuitement sur <a href="http://www.has-sante.fr">www.has-sante.fr</a>



ISBN :

HAS

Toutes les publications de l'HAS sont téléchargeables sur  
[www.has-sante.fr](http://www.has-sante.fr)