

GUIDE DE LECTURE CRITIQUE D'UN ARTICLE MEDICAL ORIGINAL (LCA)

Coordinateurs : J.L. SCHLIENGER, B. GOICHOT et N. MEYER

Faculté de Médecine de Strasbourg
Septembre 2010



PLAN

L'épreuve de lecture critique d'article original	3
Contexte, modalités, conseils	3
Objectifs pédagogiques	5
L'article original	7
Type, structure	8
Résumé	10
Grille de lecture générale	13
Technique de lecture	14
Niveau de preuve	15
Les grands types de recherche	16
Grilles de lecture spécifiques	17
Etudes épidémiologiques	17
L'essai clinique	20
La grille CONSORT	24
Le test diagnostique	27
Glossaire des termes utilisés	30
Exercices : chercher l'erreur	53
Consignes et grilles pour la rédaction du résumé	57
Articles et questions types proposés par le CNCI	
a) Dépistage de l'hépatite C	60
b) Enoxaparine/placebo	66

L'ÉPREUVE DE LECTURE CRITIQUE D'ARTICLE ORIGINAL (LCA)

1. LE CONTEXTE

L'examen classant national qui valide le 2^{ème} cycle des études médicales, il comporte une épreuve de lecture critique d'un article original. L'objectif de cette épreuve est l'incitation à lire de façon critique et à analyser le contenu d'un article. Il s'agit d'un exercice d'autoformation actuelle et future. Ce polycopié a pour but de faciliter l'apprentissage à cette lecture et son enseignement en fournissant :

- les modalités pédagogiques et docimologiques de l'épreuve
- les critères de qualité d'un article original et de son résumé
- un glossaire de méthodologie
- des grilles de lecture adaptées aux principaux types d'articles originaux
- des exemples concrets d'analyse d'article (proposés par le Conseil scientifique du CNCI)

Ces éléments devraient servir au-delà de l'apprentissage tout au long de la vie professionnelle d'un médecin quelles que soient les modalités d'exercice.

2. L'EXAMEN CLASSANT NATIONAL

L'examen Classant National (E.C.N.) comportera deux types d'épreuves destinées à explorer les capacités devant être acquises pour l'entrée du 3^{ème} cycle :

- des dossiers cliniques
- la lecture critique d'un article médical scientifique.

Les modalités actuellement retenues pour l'Examen Classant National sont les suivantes :

- trois épreuves de cas cliniques (chacune comportant trois dossiers cliniques) qui comptent pour 90 % de la note finale,
- une épreuve de lecture critique d'un article médical scientifique d'une durée de trois heures comptant pour 10 % de la note totale (arrêté du 6 décembre 2007).

3. MODALITES DE L'EPREUVE DE LECTURE CRITIQUE D'UN ARTICLE MEDICAL ORIGINAL

L'épreuve portera sur des articles scientifiques originaux, à l'exclusion des revues générales et des articles didactiques :

- traitant d'un sujet mentionné dans le programme du 2^{ème} cycle (B.O. n°31 – 30.08.2001),
- issus d'une revue médicale avec comité de lecture,
- en Français.

Le résumé et éventuellement le titre et les noms des auteurs seront supprimés.

■ Modalités de l'épreuve

Il s'agit d'une épreuve rédactionnelle : les réponses se feront avec des phrases construites, le style télégraphique ne sera pas accepté.

L'étudiant sera libre de s'organiser et de répartir son temps comme il le souhaite entre lecture et rédaction.

L'épreuve comportera deux parties : un résumé et des questions.

a) Le résumé : un résumé structuré de l'article sera demandé (la structure d'un résumé dépend bien sûr du contenu de l'article, mais le résumé comporte généralement : objectifs de l'étude, matériel et méthodes, résultats, conclusion). Le résumé ne devra pas comporter d'abréviations. Le résumé se fera dans un cadre destiné à recevoir un texte d'environ 250 mots d'une écriture moyenne.

b) Les questions

- elles seront au nombre de 6 à 10, et porteront sur les objectifs pédagogiques dont la liste est jointe en annexe (cette liste est évolutive et susceptible d'être modifiée au cours du temps)
- la longueur des réponses pourra varier selon le type de questions, et ne sera pas limitée par un cadre. Néanmoins, la longueur globale des réponses sera limitée par le cahier réservé à l'épreuve.

Modalités de correction

- Notation : résumé 30 %, questions 70 % de la note
- Compte tenu de la nature de l'épreuve, la correction sera effectuée à partir d'une grille élaborée par le jury avec l'aide d'un expert du CNCI.

4. OBJECTIFS PEDAGOGIQUES

- IDENTIFIER

Objectif 1 : L'objet d'un article médical scientifique, parmi les suivants : évaluation d'une procédure diagnostique, d'un traitement, d'un programme de dépistage, estimation d'un pronostic, enquête épidémiologique...

Objectif 2 : La " question " étudiée.

- CRITIQUER LA METHODOLOGIE

◆ Population étudiée :

Objectif 3 : Identifier les caractéristiques de la population étudiée

Objectif 4 : Analyser les modalités de sélection des sujets, critères d'inclusion et critères d'exclusion.

Objectif 5 : Analyser la technique de randomisation, le cas échéant.

Objectif 6 : Discuter la comparabilité des groupes soumis à la comparaison.

Objectif 7 : Discuter le choix des effectifs étudiés et leur cohérence dans la totalité de l'article.

◆ Méthode :

Objectif 8 : S'assurer que la méthode employée est cohérente avec le projet du travail et qu'elle est effectivement susceptible d'apporter " une " réponse à la question posée dans l'introduction.

Objectif 9 : Vérifier que les analyses statistiques (en fonction de notions élémentaires) sont cohérentes avec le projet du travail.

Objectif 10 : Vérifier le respect des règles d'éthique.

- CRITIQUER LA PRESENTATION DES RESULTATS

Objectif 11 : Analyser la présentation, la précision et la lisibilité des tableaux et des figures, leur cohérence avec le texte et leur utilité.

Objectif 12 : Vérifier la présentation des indices de dispersion (valeurs extrêmes, quantiles, écarts types) des données numériques, et celle de l'imprécision des estimations (intervalle de confiance, variance, erreur standard de la moyenne...)

- CRITIQUER L'ANALYSE DES RESULTATS ET DE LA DISCUSSION

Objectif 13 : Discuter la nature et la précision des critères de jugement des résultats.

Objectif 14 : Relever les biais qui ont été discutés. Rechercher d'autres biais éventuels non pris en compte dans la discussion et relever leurs conséquences dans l'analyse des résultats.

Objectif 15 : Vérifier la logique de la discussion et la structure. Reconnaître ce qui relève des données de la littérature et ce qui est opinion personnelle de l'auteur.

Objectif 16 : Discuter la signification statistique des résultats.

Objectif 17 : Discuter la signification clinique des résultats.

Objectif 18 : Vérifier que les résultats offrent une réponse à la question annoncée.

Objectif 19 : Vérifier que les conclusions sont justifiées par les résultats.

Objectif 20 : Indiquer le niveau de preuve de l'étude (grille de l'ANAES).

- EVALUER LES APPLICATIONS CLINIQUES

Objectif 21 : Discuter la ou les décisions médicales auxquelles peuvent conduire les résultats et la conclusion de l'article.

- CRITIQUER LA FORME DE L'ARTICLE

Objectif 22 : Identifier la structure IMRAD (Introduction, Matériel et méthode, Résultats, Discussion) et s'assurer que les divers chapitres de la structure répondent à leurs objectifs respectifs.

Objectif 23 : Faire une analyse critique de la présentation des références.

Objectif 24 : Faire une analyse critique du titre.

5. REDACTION DES QUESTIONS

- Aucun calcul nouveau ne doit être demandé.
- Un commentaire de tableaux, figures ou calculs statistiques peut être demandé.
- La forme de l'article peut être discutée ; une nouvelle rédaction du titre ou d'un paragraphe peut être demandée.
- Il ne peut être posé de questions de mémorisation sur le sujet étudié ni de questions s'écartant de l'article.
- Le résumé doit être fidèle à l'article ; il ne correspond pas à l'article tel que l'étudiant l'imagine mais tel qu'il a été conçu avec ses défauts éventuels.
- Il faut éviter les sous questions.
- Les questions ne sont pas posées dans un langage de spécialistes. Elles doivent être comprises simplement par tout étudiant de deuxième cycle n'ayant pas reçu de formation particulière.
- Si une définition pose problème, c'est-à-dire si elle peut faire l'objet de discussions, la définition retenue pour cette épreuve doit être précisée dans l'intitulé.

6. MODALITES DE CORRECTION

- Le résumé ne doit pas dépasser 250 mots. L'étudiant indique le nombre de mots. Une pénalité sera décidée par le jury en cas de dépassement, proportionnelle au degré de dépassement.
- Une grille de réponses est fournie par le conseil scientifique du CNCI pour les questions portant sur l'article (sur 100 points non divisibles) et pour le résumé (sur 50 points non divisibles).
- Des points négatifs peuvent être décidés par le jury.
- La correction est une double correction comme les dossiers.

L'ARTICLE ORIGINAL

1. LE TYPE D'ARTICLE

La bonne connaissance des principes de rédaction d'un article original fournit de précieux repères pour sa lecture critique.

Il existe **différents types d'articles** :

- original
- mise au point ou revue de la littérature
- méta-analyse
- article préliminaire
- article didactique
- lettre à la rédaction

L'article original rapporte les résultats d'une étude visant à confirmer ou infirmer une hypothèse de travail nouvelle. On distingue différents types d'articles originaux selon leurs objectifs.

TYPE D'ARTICLE ORIGINAL	OBJECTIF	SCHEMA EXPERIMENTAL
Essai thérapeutique	Evaluer l'efficacité d'un traitement	Essai comparatif randomisé, si possible en double insu
Test diagnostique	Démontrer la validité et la fiabilité d'un test	Comparaison d'un nouveau test à un test de référence
Recherche d'association et facteurs de risque	Déterminer si un facteur est responsable d'un événement	Etude de cohorte ou étude cas-témoin
Etude pronostique	Etudier le devenir des malades	Etude de cohorte
Etude descriptive	Caractériser une population	Transversale ou série de cas

2. STRUCTURE D'UN ARTICLE ORIGINAL

Un article original est structuré autour des **règles IMRAD**
(**I**ntroduction, **M**éthodes, **R**ésultats and **D**iscussion)

La connaissance des recommandations facilite la lecture critique d'un texte.

A. LE TITRE : bref, clair, sans abréviation avec le maximum d'informations sur l'objet de l'article avec un minimum de mots. Reflet exact de l'article, attractif, sélectif mais non provocateur.

B. L'INTRODUCTION : doit permettre de **comprendre pourquoi le travail a été entrepris** et quel était son objectif

2 parties :

1. Objectif informatif : apporter au lecteur les éléments nécessaires et suffisants à la compréhension du travail (passerelle entre les connaissances de l'auteur et celles du lecteur).

Description du contexte qui a présidé à la mise en œuvre de l'étude.

2. Préciser le but du travail la question posée et l'objectif. L'hypothèse doit être clairement exposée. Chaque affirmation doit être référencée.

C. PATIENTS ET METHODES

Suffisamment précis pour qu'un lecteur puisse reproduire ou vérifier le travail. Le nœud de la rigueur scientifique et de la crédibilité.

Ne comporte ni résultats ni commentaires.

Cette section décrit comment le travail a été réalisé en fournissant des informations précises, rigoureuses, devant permettre de reproduire le travail en totalité.

1^{er} OBJECTIF : les patients de l'étude (le " matériel ")

- contient tous les détails nécessaires à l'interprétation des résultats
- doit permettre d'identifier les biais éventuels
- précise le critère principal de jugement et les critères accessoires.

La **description comporte :**

- les critères d'inclusion ou de non-inclusion, ils doivent être adaptés au but de l'essai (âge, sexe, CSP¹, origine ethnique, symptomatologie, données morphologiques, groupes ou sous-groupes)
- le schéma expérimental : série consécutive ou non ; travail prospectif ou rétrospectif, randomisé ou non, ouvert ou en insu, primo traitement ou non
- les éléments ayant permis la construction d'un échantillon ou d'une population, calcul de la taille de l'échantillon
- traitements étudiés ; le traitement de référence est-il utilisé suivant les modalités habituelles ?
- Accord du CPP, consentement des patients.

¹ CSP : Catégorie socio-professionnelle

2ème OBJECTIF : ce que l'on cherche à évaluer

- Action d'un médicament, résultats d'une intervention, valeur diagnostique d'un examen.
- Le critère de jugement principal et éventuellement un ou des critères secondaires est (sont) sont précisé(s) ainsi que les moyens permettant de les évaluer.

3ème OBJECTIF : appréciation des résultats

- Valeurs normales retenues
- Tests et méthodes statistiques
- Précision : comment chiffrer un amaigrissement, comment quantifier une diarrhée
- Indiquer le nombre de malades exclus de l'analyse (en donnant la raison) ou perdus de vue

D. RESULTATS

Ce chapitre " résultats " ne doit contenir que les résultats se référant directement aux critères de jugement.

1. Tous les résultats et rien que les résultats, y compris les résultats négatifs dans la mesure où ils apportent une information utile.

En cas d'essai thérapeutique la répartition des groupes et leur comparabilité à l'inclusion sont mentionnées, de même que le nombre et la cause des arrêts prématurés et des sorties d'essai, la répartition des perdus de vue et les effets secondaires.

2. Les erreurs possibles de l'auteur

- donner les résultats qui ne sont pas cohérents avec le but du travail (confusion)
- faire des commentaires

3. Tableaux et figures

Ils fournissent le *maximum d'information* sous une forme synthétique et claire. Ils évitent les répétitions entre l'exposé des résultats et leur rappel comme base de discussion dans le chapitre " discussion ".

Ils doivent être *informatifs* par eux-mêmes (légende, titre...)

Ils doivent être appelés dans le texte.

E. DISCUSSION

1. Précise si le but du travail exposé à la fin de l'introduction a été atteint ou non

2. Juge de la qualité et de la validité des résultats : la discussion critique et objective porte sur chacun des chapitres de l'article. Commentaire sur les biais.

3. Compare les résultats observés à ceux d'autres auteurs en cherchant à expliquer les différences et en soulignant l'apport original.

Les erreurs

- Répétition de ce qui a été dit dans l'introduction
- l'inexactitude des citations

- l'usage des temps : au passé tout ce qui relate un fait (résultat) ; au présent ce qui est opinion d'auteur et ce qui correspond à des notions bien établies

F. REFERENCES

1. **Elles justifient tout fait énoncé** : principe fondamental de la démarche scientifique.
2. **Elles soutiennent les faits ayant amené les auteurs à formuler l'objectif du travail** (introduction), renvoient aux méthodes déjà décrites (matériel et méthodes) et argumentent les résultats critiqués dans la discussion
3. **Sélectionnées de façon pertinente**, elles doivent être appelées dans le texte

3. RESUME

Un bon résumé informatif doit pouvoir être lu indépendamment de l'article par un lecteur non spécialiste. Strictement fidèle au texte, il doit inciter à le lire.

A. PRINCIPES GENERAUX

Informatif il ne doit pas dépasser **250 mots**, ne comporter ni tableau, ni abréviation, ni référence bibliographique. Il est rédigé au passé (phrases construites avec un verbe) sauf pour l'introduction et la conclusion. Il répond aux quatre questions fondamentales :

- pourquoi ce travail a été fait ?
- comment ?
- qu'est-ce qui a été trouvé ?
- quelles conclusions ou quelle généralisation possible et autorisée
- sa construction reprend la structure IMRAD : introduction, matériel et méthode, résultats, discussion (limitée habituellement à une phrase de conclusion)
- la première phrase exprime l'idée directrice (sans répéter le titre)
- la seconde énonce la méthode de travail
- les erreurs à ne pas commettre :
 - o le résumé indicatif (au lieu d'informatif), réservé à la revue générale
 - o éviter les abréviations
 - o informations trop peu détaillées pour comprendre l'article

B. POINTS DE REPERES POUR LA REDACTION DU RESUME

Strictement limité à 250 mots, écrit au passé il ne comporte ni tableau, ni figure. Il est recommandé de mettre des inter-titres (introduction, méthodes, etc...) et d'indiquer un titre qui ne sont pas décomptés. Ecrit de façon très lisible en respectant la grille, il fait l'objet d'un "comptage" par l'étudiant en évitant les abréviations et selon les modalités indiquées en annexe. Le nombre de mots doit être indiqué. (Voir la grille de rédaction et les modalités de comptage des mots en annexe)

INTRODUCTION contexte : 1 à 2 phrases pour le rationnel de l'étude
Objectif : quel est l'objectif principal ?

MATERIEL ET METHODE

Schéma expérimental :

- essai thérapeutique contrôlé, randomisé, en double insu ?
- test diagnostique : quel est le critère de référence ?
- étude pronostique : étude de cohorte ?
- étude de causalité : essai contrôlé, randomisé ? cohorte ? cas témoins ?
- étude descriptive : transversale ? série de cas ?

Lieu de l'étude

- communauté, soins primaires, centre spécialisé, ambulatoire ou en hospitalisation

Patients

- critères d'éligibilité (inclusion, non inclusion)
- nombre de patients inclus

Intervention

- description et durée, nature de l'intervention

Critères de jugement

- critère principal et critères accessoires retenus avant le début de l'étude.

RESULTATS

- principaux résultats – moyennes, intervalles de confiance, signification statistique.
Indiquer tous les résultats concernant le critères principal. Eviter les résultats
" distractifs.

CONCLUSIONS

Elles portent strictement sur les résultats, éviter les spéculations et les généralisations hâtives . Il n'y a pas de discussion des résultats. Autant que possible mettre en perspective et d'extrapoler la portée des résultats).

MOTS CLES (pour l'ECN)

3 à 10 mots répertoriés dans l'Index Medicus.

Les " trucs " de rédaction du résumé (ECN)

- Ni délayage, ni style télégraphique : concision et précision.
- Ecrire les têtes de paragraphe pour débiter.
- Relever les phrases ou les fragments les plus significatifs dans l'article.
- Identifier l'objectif principal et la conclusion.
- Les mots-clés du paragraphie " méthodes "
- Les résultats principaux.

- Formuler l'hypothèse dans la première phrase du résumé.
- Respecter la cohérence.
- Ecrire au " passé " (le présent peut être employé dans la conclusion).
- Un " résumé " des résultats principaux est souvent fourni par les auteurs au début du paragraphe discussion (veillez à valider sa conformité avec les résultats).
- Ne rien écrire qui ne figure pas explicitement dans l'article.
- Exposer les résultats strictement comme dans le texte.
- Eviter les travers " les auteurs font... les auteurs analyses, proposent, estiment, etc... "
- Relire afin de vérifier que le résumé ne contient que l'information essentielle rendue sous une forme " lisible ".

GRILLE DE LECTURE GENERALE

(Audet N. Ledère H., *Pédagogie Médicale*, 2001, 2, 206-212)

Une grille de lecture permet

- La compréhension globale d'une étude.
- L'évaluation de sa qualité méthodologique.
- L'évaluation de l'importance du résultat pour la pratique clinique.

1. LIRE DE FACON CRITIQUE : L'INVENTAIRE

A) DETERMINER LA VALEUR SCIENTIFIQUE D'UN ARTICLE

- **pertinence de l'article** par rapport à son questionnement
- évaluer la **justesse des données**
 - . vérifier l'authenticité des faits
 - . s'assurer de l'exhaustivité des données
 - . distinguer un fait d'une opinion
 - . détecter les biais (erreurs systématiques) : évaluer si la constitution de l'échantillon permet la généralisation des résultats et si la taille est suffisante pour être pertinente
 - . dépister les relations de cause à effet erronées
 - . reconnaître les affirmations non valides
- évaluer l'**interprétation des données**
 - . critiquer le choix des tests statistiques
 - . déterminer le niveau de preuve proposé par le plan de recherche
 - . critiquer le choix du plan de recherche
- identifier les **généralisations fautives**

B) APRES LA LECTURE, REALISER UNE SYNTHESE portant sur l'utilité clinique des résultats

- en général
- pour sa propre pratique
- capacité à induire un changement de comportement professionnel
- utilité en tant qu'outil d'autoformation en cours de lecture (métacognition)

2. LA TECHNIQUE DE LECTURE D'UN ARTICLE : cadre de référence

CHOISIR CE QU'ON LIT

- motivation** : préoccupation fondamentale ou clinique
- posséder une certaine maîtrise du sujet
 - disposer d'une stratégie efficace de recherche d'articles
 - banques de données, Internet
 - bibliothèque
 - expert
 - archivage d'articles

DECIDER DE LA LECTURE

- être concerné par le sujet
- être familier avec le contenu médical et les concepts de biostatistique et d'épidémiologie clinique utilisés.

IDENTIFIER LA PERTINENCE

- a. **Comprendre les concepts méthodologiques utilisés**
- b. **Evaluer la justesse des données**
 - distinguer un fait d'un commentaire
(IMRAD – introduction, méthodes, résultats and discussion)
 - détecter les biais
 - vérifier l'exhaustivité des données
 - s'assurer de la qualité et de la pertinence des tableaux et figures
- c. **Evaluer l'interprétation des données**
 - niveau de preuve
 - plan de recherche
 - tests statistiques utilisés
 - généralisation des résultats en accord avec les limites permises par la méthodologie pour la population de référence

Cette étude est ou n'est pas valide sur le plan scientifique.

Cette étude est ou n'est pas utile pour la pratique médicale.

Elle doit être confrontée à la grille de graduation des recommandations.

GUIDE D'ANALYSE DE LA LITTÉRATURE : NIVEAU DE PREUVE ET GRADATION DES RECOMMANDATIONS

Agence Nationale d'accréditation et d'évaluation en santé (ANAES/HAS)

Niveau de preuve des études	Grade des recommandations
Niveau 1 <ul style="list-style-type: none">- essais comparatifs randomisés de forte puissance- méta-analyse d'essais comparatifs randomisés- analyse de décision basée sur des études bien menées	A Preuve scientifique établie
Niveau 2 <ul style="list-style-type: none">- essais comparatifs randomisés de faible puissance- études comparatives non randomisées bien menées- études de cohortes	B Présomption scientifique
Niveau 3 <ul style="list-style-type: none">- études cas-témoin Niveau 4 <ul style="list-style-type: none">- études comparatives comportant des biais importants- études rétrospectives- séries de cas- études épidémiologiques descriptives (transversale, longitudinale)	C Faible niveau de preuve scientifique

Indépendamment de la méthode utilisée (parmi les méthodes présentées à la page précédente), un article se caractérise par l'objet principal de son étude, c'est-à-dire par le type de recherche qui y est présenté.

Les grands types de recherche

La plupart des études s'inscrivent dans l'une ou l'autre des catégories suivantes :

- **Thérapeutique** : mesurer l'efficacité d'un médicament, d'une intervention chirurgicale, d'une nouvelle méthode d'éducation des patients ou de toute autre procédure. Le protocole de choix est l'essai comparatif randomisé.
- **Diagnostic** : montrer si un nouveau test diagnostique est valide (peut-on lui faire confiance ?) et fiable (fournit-il les mêmes résultats lorsqu'on le répète ?). Le protocole de choix est l'enquête transversale (*cross sectional survey*) où sont mis en œuvre simultanément le nouveau test et l'examen de référence (*gold standard*).
- **Dépistage** : estimer la valeur d'un test qui peut être appliqué à une vaste population et qui détecte une maladie à un stade présymptomatique. La méthode de choix est l'enquête transversale (*cross sectional survey*).
- **Pronostic** : prévoir comment évoluera une maladie diagnostiquée chez un individu à un stage précoce. La méthodologie de choix est l'étude de cohortes longitudinale.
- **Etiologie** : démontrer si un agent présumé pathogène, une pollution environnemental par exemple, est effectivement impliqué dans la survenue d'une maladie. Le protocole de choix est soit l'étude de cohortes, soit l'étude cas-témoins, selon la fréquence de la maladie mais les séries de cas peuvent également apporter des informations cruciales.

A chacun de ces grands types de recherche correspond un type d'article pouvant être décortiqué avec une grille de lecture générale mais aussi avec une grille qui lui est spécifique. On ne lit pas un article relatif à un essai thérapeutique comme un article consacré au développement d'un test diagnostique. Des grilles de lectures spécifiques vous sont donc proposées ici.

GRILLES DE LECTURES SPECIFIQUES

1. Etudes épidémiologiques*

Par niveau de preuves décroissant on distingue les études de cohorte, les études descriptives et transversales de population et les études cas témoins.

- L'étude de cohorte consiste à comparer l'incidence d'événements bien définis dans deux populations suivies pendant un temps donné, l'une exposée et l'autre non exposée. Elle permet le calcul du risque relatif. Le nombre de sujets nécessaires dépend de l'incidence attendue de l'événement dans la population et l'augmentation estimée du risque (ex tabac et cancer, dyslipidémie et accidents coronaires). Les études de cohorte peuvent contribuer à préciser les facteurs prédictifs ou pronostiques. L'étude de cohorte n'est pas réalisable pour des événements rares ou retardés nécessitant une population très importante ou une durée de suivi très prolongée.
- L'étude pronostique : c'est un cas particulier d'étude de cohorte qui s'adresse à une série de sujets déjà porteurs de la maladie à l'étude et pour lesquels on étudie le rôle pronostique (risque de rechute, de décès, d'aggravation) de différents facteurs.
- L'étude cas-témoin : elle compare les cas à des témoins en recherchant dans les deux groupes l'exposition à un facteur causal. Elle permet de calculer " l'Odds Ratio " ou rapport des cotes. Le groupe témoin doit idéalement avoir les mêmes caractéristiques que les cas. Elle est utile lorsque l'étude de cohorte n'est pas possible ou lorsqu'une étude expérimentale n'est pas possible.
- L'étude transversale descriptive d'une population permet d'établir des relations entre divers paramètres sans pouvoir affirmer qu'il existe une relation causale.
- La série de cas.

L'information existe-t-elle pour chacune de ces 8 questions ?	La façon d'aborder la question est-elle correcte ?	Impact sur les conclusions
1.Objectif - Pronostic – évolution - Impact d'une intervention - Etiologie – causalité	- Y a-t-il une hypothèse ? - L'objectif principal est-il clairement défini ?	
2. Type d'étude - Essai contrôlé randomisé - Etude de cohorte - Etude cas-témoins - Etude transversale - Rapport de cas – série de cas	- Le type de l'étude est-il approprié à la question posée ? - S'agit-il du type d'étude apportant le niveau de preuve le plus élevé	- Si non, les résultats de l'étude sont-ils totalement inutiles ?

* Grille proposée par M. Velten et Mme C. Simon

<p>3. Facteur(s) étudié(s) - Exposition - Intervention - Test diagnostique</p>	<p>Sont-ils bien décrits ? Comment sont-ils mesurés ? - Même méthode de mesure chez tous les sujets ? dans tous les groupes ? - Méthode à l'aveugle ? Y a-t-il une comparaison indépendante avec l'étalon ?</p>	<p>Sinon ce biais de mesure menace-t-il la validité de l'étude ? <i>Idem</i> Sinon ce biais menace-t-il la validité de l'étude ?</p>
<p>4. Critère(s) de jugement</p>	<p>Le critère de jugement est-il unique ? Si non, a-t-on défini un critère principal Comment sont-ils mesurés ? - <i>Idem</i> question 3 Tous les critères de jugement pertinents sont-ils évalués ? En cas de critères multiples, les résultats sont-ils convergents ?</p>	<p>- <i>Idem</i> question 3 Si non, ceux qui ont été oubliés sont-ils importants ?</p>
<p>5. Population source et sujets étudiés</p>	<ul style="list-style-type: none"> - Les critères d'inclusion et d'exclusion sont-ils précisés et pertinents ? - Y a-t-il des perdus de vue ? Si oui, leur nombre et leur répartition sont-ils précisés ? - Y a-t-il des non-réponses ? Comment ont-elles été prises en compte ? - A-t-on vérifié la robustesse des conclusions selon les hypothèses faites pour les non-réponses ? - Y a-t-il randomisation ? A-t-elle été effectuée correctement ? Les effectifs permettant de vérifier les caractéristiques des sujets inclus et leur bonne répartition sont-ils présentés ? - Les groupes diffèrent-ils par des caractéristiques autres que les facteurs étudiés ? - Quelle est la proportion de sujets atteignant la fin du suivi ? - Les perdus de vue sont-ils similaires aux sujets encore suivis ? 	<p>- Si des biais sont à craindre, pour chacune des questions précédentes, cela menace-t-il :</p> <ul style="list-style-type: none"> - la validité interne ? - la validité externe ?
<p>6. Facteurs de confusion potentiels et biais</p>	<ul style="list-style-type: none"> - Sont-ils tous envisagés ? - Sont-ils bien contrôlés ? - Biais d'information (de mesure, cf 3.) ? Erreurs différentielles ? Erreurs non différentielles ? - Biais de sélection ? 	<ul style="list-style-type: none"> - Si non, cela invalide-t-il l'étude ? - Situation rendant l'étude inexploitable, si importantes - Conclusion possible uniquement si résultat significatif - L'interprétation est-elle encore possible ?

<p>7. Résultats</p> <p>Type d'analyse utilisée ? Intervalles de confiance ?</p> <p>Tests statistiques</p> <p>- Si résultats positifs - Si résultats négatifs</p> <p>Force de l'association</p> <p>Cohérence externe</p>	<p>Est-elle adaptée ? A-t-on privilégié la présentation des estimations accompagnées d'un intervalle de confiance ?</p> <p>- Réalisés uniquement pour les hypothèses indiquées au départ ? - Respect des conditions d'application ?</p> <p>- Cliniquement intéressant ? - Taille de l'échantillon suffisante (puissance) ?</p> <p>- Est-elle indiquée ? - L'indicateur choisi est-il pertinent ? - Peut-on vérifier les résultats à partir des données présentées ?</p> <p>Les résultats sont-ils cohérents - avec ceux obtenus par d'autres études épidémiologiques ? - avec ceux obtenus dans d'autres disciplines, par d'autres méthodes ?</p>	<p>- Si non, peut-on les calculer à partir des données présentées ?</p> <p>- Le risque global d'erreur de première espèce est-il conservé ?</p> <p>- Si non, l'étude est-elle utile ? - Si insuffisant, l'étude est-elle utile ou non concluante ?</p>
<p>8. Conclusions des auteurs ?</p> <p>Réponses aux questions ? Vérification de l'hypothèse ? Objectif atteint ?</p>	<p>- Les conclusions répondent-elles à l'objectif ?</p>	<p>Les résultats sont-ils acceptables appliqués à la population source (validité) ?</p> <p>Les résultats peuvent-ils être extrapolés aux situations qui vous intéressent pour votre propre pratique (applicabilité) ?</p>

2. L'essai clinique (J.L. IMBS)

La lecture critique d'une publication d'essai clinique procède en trois temps :

- identification de l'objectif de l'étude,
- analyse méthodologique,
- pertinence des résultats pour la pratique clinique.

1. OBJECTIF

L'essai ne peut répondre qu'à une seule question : cet **objectif (principal)** doit être clairement décrit. Sa définition fera choisir un **critère de jugement principal** qui sera, entre autre, utilisé pour calculer le **nombre de sujets nécessaire**. Des **objectifs secondaires** peuvent y être associés : ainsi une analyse de la sécurité d'emploi du médicament (la tolérance) peut s'ajouter à un objectif principal ciblé sur l'évaluation de l'efficacité.

Il est souhaitable que l'utilité thérapeutique de la question posée soit explicitée, ou du moins située dans le contexte actuel (ces données étant référencées dans la bibliographie).

Schématiquement, il faut vérifier que :

- le but de l'essai est défini,
- de façon compréhensible ;
- la réponse à la question posée pourrait avoir une application clinique.

2. METHODOLOGIE

Un essai **prospectif, contrôlé** (vs placebo ou vs un **traitement de référence** reconnu), **randomisé** selon un **plan expérimental** adéquat mené en **double-insu** sur un **échantillon** parfaitement défini dont l'effectif aura été **calculé** selon la **puissance** choisie, utilisant un **critère de jugement** adapté à l'objectif, offrira un **niveau de preuve** optimal.

Ces paramètres seront considérés selon 3 questions (a, b, c)

a) *Comment les réponses aux traitements testées vont-elles être mesurées ?*

- Selon un plan expérimental correctement choisi (groupes parallèles, chassé-croisé, factoriel, séquentiel)
- Le critère d'évaluation (principal) répond bien à l'objectif de l'étude. Le choix des horaires de sa mesure, la sensibilité et la spécificité de l'instrument de mesure, éventuellement la variabilité inter-individuelle (paramètre utile pour le calcul du nombre de sujets nécessaire) sont connus. Autrement dit, il faut vérifier que : un critère principal est défini pour mesurer l'efficacité ; ce critère principal a une signification médicale claire ; il est adapté au but de l'essai.
Chaque fois que possible c'est un critère quantitatif qui sera choisi (plutôt que qualitatif).
- L'échantillon de la population étudiée est parfaitement défini par des **critères d'inclusion et de non-inclusion**. Il faut se souvenir que des critères d'inclusion étroits permettront d'obtenir une homogénéité des patients inclus favorable à la détection d'une significativité statistique mais que les résultats obtenus sur cet échantillon très particulier pourront ne pas être extrapolables à la généralité des patients souffrant de l'affection considérée.
- Le choix du **traitement de référence** (ou traitement " contrôle ") répond aux données actuelles du traitement de l'affection considérée, qu'il s'agisse de l'emploi d'un **placebo** ou d'une spécialité pharmaceutique qui devra alors être utilisée à la posologie et pour la durée reconnues dans l'indication faisant l'objet de l'étude.

b) *Comment la comparabilité des groupes est-elle satisfaite ?*

- La répartition des traitements est faite par **tirage au sort**. La méthode utilisée est correcte (le plus souvent : table de nombre au hasard) S'il existe une stratification, elle est justifiée sur le plan clinique et ne multiplie pas les sous-groupes de façon excessive et préjudiciable à la puissance de l'étude.
- Le **double-insu** (éventuellement **simple insu**) est bien assuré, au besoin avec la participation d'observateurs externes (lorsque la mesure d'un critère de jugement risque de lever l'aveugle). La méthode du " double-placebo " est éventuellement utilisée, à bon escient.
- Le nombre de patients sortis de l'étude ou perdus de vue n'est pas différent selon les groupes (ou les périodes).

c) *Comment la puissance de l'essai est-elle assurée ?*

- Le **nombre de sujet nécessaire** a été calculé. Les risques α (1^{ère} espèce) et β (2^{ème} espèce) acceptés sont indiqués. Cet effectif est atteint.
- Le nombre de sorties d'étude (< 10 à 15 % des inclus), de **perdus de vue** (< quelques %), ou de valeurs manquantes, n'est pas invalidant.
- L'**observance** des traitements comparés a été mesurée. Les moyens de favoriser cette observance ont été mis en œuvre.
- En cas d'**essai multicentrique**, le choix des centres participants est raisonnable et des efforts ont été faits pour standardiser au niveau de chacun d'eux les procédures de l'étude. Un tirage au sort séparé, équilibré par centre, a été fait.
- Le recueil des données et leur saisie ont été organisés de façon à réduire les erreurs de transcription.

3. ANALYSE DES RESULTATS

- a) La comparabilité des groupes expérimentaux à l'inclusion (au début de l'étude) est vérifiée. Le choix des critères de comparabilité est raisonnable.
- b) Des tableaux rendant compte des effectifs de sujets inclus, sortis d'étude, perdus de vue, avec les effectifs utilisés pour une analyse en **intention de traiter** ou en "**per protocole**" sont fournis. Les chiffres qu'ils mentionnent sont concordants, en particulier avec le texte de la publication.
- c) Le test statistique utilisé pour analyser la différence entre les valeurs du critère de jugement principal recueillies pour chaque groupe expérimental est bien choisi. L'analyse statistique porte sur la totalité des patients inclus dans l'essai (analyse en intention de traiter). Il existe une différence avec un $p < 0.05$. Ce résultat est retrouvé par l'analyse per protocole. Cette différence statistiquement significative est cliniquement intéressante.
- d) Les résultats d'éventuelles analyses en sous-groupe ne sont donnés qu'à titre informatif.
- e) Les résultats de l'analyse d'éventuels critères de jugement secondaires sont cohérents.

4. CONCLUSION

Elle reprend les faits démontrés sans donner d'opinion ou formuler de nouvelle(s) hypothèse(s). Elle indique si le résultat obtenu peut être extrapolé à l'ensemble des formes cliniques de l'affection traitée ou doit rester limité au type d'échantillon recruté pour l'étude. Elle souligne l'éventuel bénéfice thérapeutique pratique qui résulte de l'essai.

* Les notions explicitées dans le glossaire figurent en **caractères gras** lors de leur première citation dans cette grille.

LES DIX COMMANDEMENTS DE L'ESSAI THERAPEUTIQUE

(Grille de lecture d'un essai thérapeutique ou de prévention)

1. Essai contrôlé ?
2. Etude randomisée ?
3. Essai en double aveugle ?
4. Nombre de sujets nécessaires calculé.
5. Définition claire de la population.
6. Un critère principal de jugement unique et clinique pertinent.
7. Analyse en intention de traiter.
8. Une analyse en fin d'essai sur la totalité des patients.
9. Différence statistiquement significative.
10. Une différence cliniquement significative.

ANALYSE D'UN ARTICLE D'ESSAI THERAPEUTIQUE A L'AIDE DE LA GRILLE CONSORT.

La grille CONSORT (Consolidated Standards of Reporting Trials) est un outil-guide international destiné à évaluer la qualité de la rédaction d'un essai contrôlé randomisé.

1. Titre et résumé.
2. Introduction : rationnel.
3. Critères d'inclusion, d'exclusion.
4. Description des traitements alloués.
5. Objectifs spécifiques de l'étude et hypothèses.
6. Critères de jugement principal et secondaire.
7. Détermination des effectifs nécessaires.
8. Méthode de randomisation.
9. Méthode illustrant les unités thérapeutiques après randomisation.
10. Définition des investigateurs responsables de l'inclusion, de la randomisation et de la distribution des unités thérapeutiques.
11. Simple aveugle ou double aveugle.
12. Méthodes statistiques utilisées pour le critère de jugement principal, analyse par sous-groupes ?
Ajustements ?
13. Diagramme décrivant le nombre de patients à chaque stade de l'étude (flow chart) : screening, inclusion, randomisation, allocation des différents traitements, suivi de l'étude (perdu de vue, arrêt de traitement), nombre sur lequel a porté l'analyse.
14. Date de la période de recrutement et follow-up.
15. Caractéristiques cliniques et démographiques au début de l'étude des deux groupes comparés.
16. Nombre de participants dans chaque groupe et nombre de patients sur lequel porte l'analyse (analyse en intention de traiter).
17. Résultats précis avec intervalle de confiance à 95 % sur le critère de jugement principal et critères de jugement secondaires.
18. Analyses par sous-groupes, analyses après ajustement, statistiques.
19. Description des effets secondaires dans chaque groupe.
20. Interprétation des résultats, discussion des biais ou imprécisions et des dangers éventuels.
21. Applicabilité clinique des résultats de l'étude, les résultats de l'étude sont-ils extrapolables à la population générale (validité externe).
22. Interprétation générale des résultats et recommandations.

QUESTIONS COMPLEMENTAIRES A LA GRILLE CONSORT

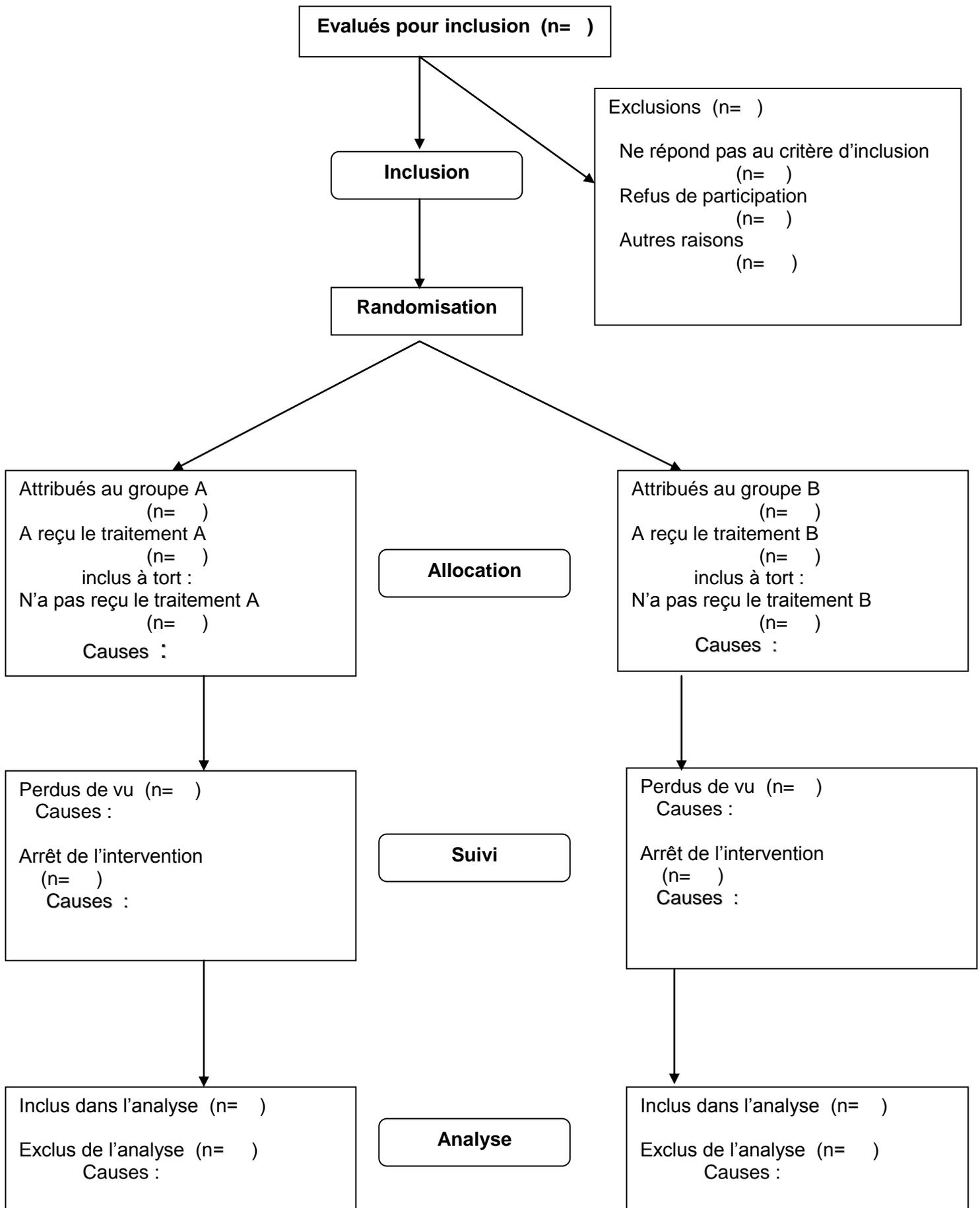
- Pertinence de la question : importance, originalité.
- L'hypothèse testée est-elle clairement exprimée ?
- Quelle est la comparabilité exprimée ?
- Quelle est la comparabilité des groupes ?
- Les facteurs pronostiques susceptibles d'intervenir ont-ils été pris en compte (stratification dans la randomisation, ajustements lors de l'analyse statistique).
- Cause des arrêts prématurés de traitement.
- Pertinence clinique d'une différence observée.
- Quel est le nombre de sujets à traiter pour éviter un événement ?
- L'étude est-elle éthique ?
- Les conclusions de l'étude sont-elles en rapport avec l'hypothèse de travail et concernent-elles le critère de jugement principal ?

LECTURES CONSEILLEES

- Dossier " Essais cliniques " paru dans la Revue du Praticien, 2000, 50, p. 827 à 869, et en particulier : BERGMANN J.F. et CHASSANY O. Lire un compte-rendu d'essai clinique, p. 838 à 845.
- BOUVENOT G., VILLANI P., AMBROSI P. Lecture critique de la publication d'un essai clinique. Presse Med., 2002, 31 : 1061-8.
- Dossier " La vie du médicament " paru dans la Revue du Praticien, 2002, 52 : 480-515

DIAGRAMME DE FLUX (EN ANGLAIS : FLOWCHART)

Le diagramme de flux décrit sous forme d'organigramme le recrutement, la randomisation et le suivi de l'ensemble des patients lors de l'étude. Il permet de connaître rapidement le nombre d'inclusions, l'équilibre des groupes et l'éventuelle perte de cet équilibre au long de l'essai. Il permet donc de connaître à chaque étape de l'étude le devenir des sujets ayant participé à l'étude. Le diagramme de flux donné ci-dessous est le diagramme de la grille CONSORT.



3. Le test diagnostique*

NB : par " test ", on entend toute information dont le résultat peut être utilisé dans une démarche décisionnelle, le plus souvent dans un but diagnostique. Ce peut être un examen biologique ou d'imagerie, un signe clinique, une procédure ou une stratégie (c'est à dire la combinaison de plusieurs tests)

1. Identifier ce qui est évalué : examen biologique ou autre, signe clinique, stratégie....
 - Le test évalué est-il suffisamment bien décrit pour être reproduit ?

2. Identifier l'examen de référence (" gold standard ")
 - Le test étudié doit être comparé à l'examen le plus performant disponible pour affirmer le diagnostic. Le choix de cet examen de référence est capital, il doit être justifié dans le texte avec des références adéquates. Lorsque cet examen n'existe pas, cela peut être une vérification autopsique ou chirurgicale, plus rarement une évolution clinique qui confirme a posteriori le diagnostic.
 - Tous les patients étudiés doivent avoir les deux tests, celui qui est évalué et celui de référence, et l'interprétation de chacun de ces tests doit être faite de façon indépendante, en " aveugle ".
 - La maladie et l'absence de maladie sont elles bien définies ?

3. La population étudiée est elle correctement décrite ?
 - Quelle est la fréquence de la maladie dans cette population ? Cette fréquence sera elle identique dans les populations auxquelles les auteurs destinent ce test ?
 - S'il y a des stades de gravité différente pour cette maladie, sont-ils mesurés ?

4. La reproductibilité du test est-elle analysée (variabilité intra/inter observateur) ? L'interprétation du test dépend elle de l'expérience d'un examinateur (par ex. examen radiologique) ?

5. Comment sont exprimées les qualités diagnostiques du test : sensibilité, spécificité, valeurs prédictives, rapports de vraisemblances (voir glossaire) ?

6. Les intervalles de confiance sont-ils fournis ?

7. L' " applicabilité " du test est elle discutée en tenant compte : de la fréquence de la maladie dans d'autres populations ? de la faisabilité technique ? d'une évaluation économique ? Ce test a t-il été placé dans le contexte des autres examens utilisés et de leur séquence ?

* Grille proposée par B. Goichot et S. Rohr

Rappels utiles pour le "test diagnostique"

	Malade	Non malade	Total
Test positif	VP	FP	VP+FP
Test négatif	FN	VN	FN+VN
Total	VP+FN	FP+VN	n

Lecture verticale : $Se = VP/(VP+FN)$ $Sp = VN/(VN+FP)$

Lecture horizontale : $VPP = VP/(VP+FP)$ $VPN = VN/(VN+FN)$

Sensibilité (Se) = probabilité d'avoir un test positif quand on est malade

VPP = probabilité d'être malade quand on a un test positif

Spécificité = probabilité d'avoir un test négatif quand on n'est pas malade

VPN = probabilité de ne pas être malade quand on a un test négatif

Sensibilité et spécificité décrivent les qualités *intrinsèques* du test.

Les valeurs prédictives dépendent des qualités intrinsèques du test et de la *probabilité pré-test*, c'est à dire de la prévalence de la maladie dans la population étudiée. Elles décrivent l'utilité diagnostique du test dans *une population donnée*.

Les *rapports de vraisemblance* (LR likelihood ratio) décrivent l'apport d'un test au diagnostic :

Le rapport de vraisemblance positif (L) d'un test est la vraisemblance de la présence de la maladie lorsque le test est positif. Il varie de 0 à l'infini. Si $LR = 1$, cela signifie que la probabilité après avoir fait le test est identique à la probabilité pré-test, donc que le test n'apporte rien au diagnostic. Plus il est élevé, plus le " gain diagnostique " est important.

Le rapport de vraisemblance négatif (λ) est la vraisemblance d'être malade lorsque le test est négatif. Il varie entre 0 et 1. Plus il est proche de 0, plus il permet d'exclure le diagnostic.

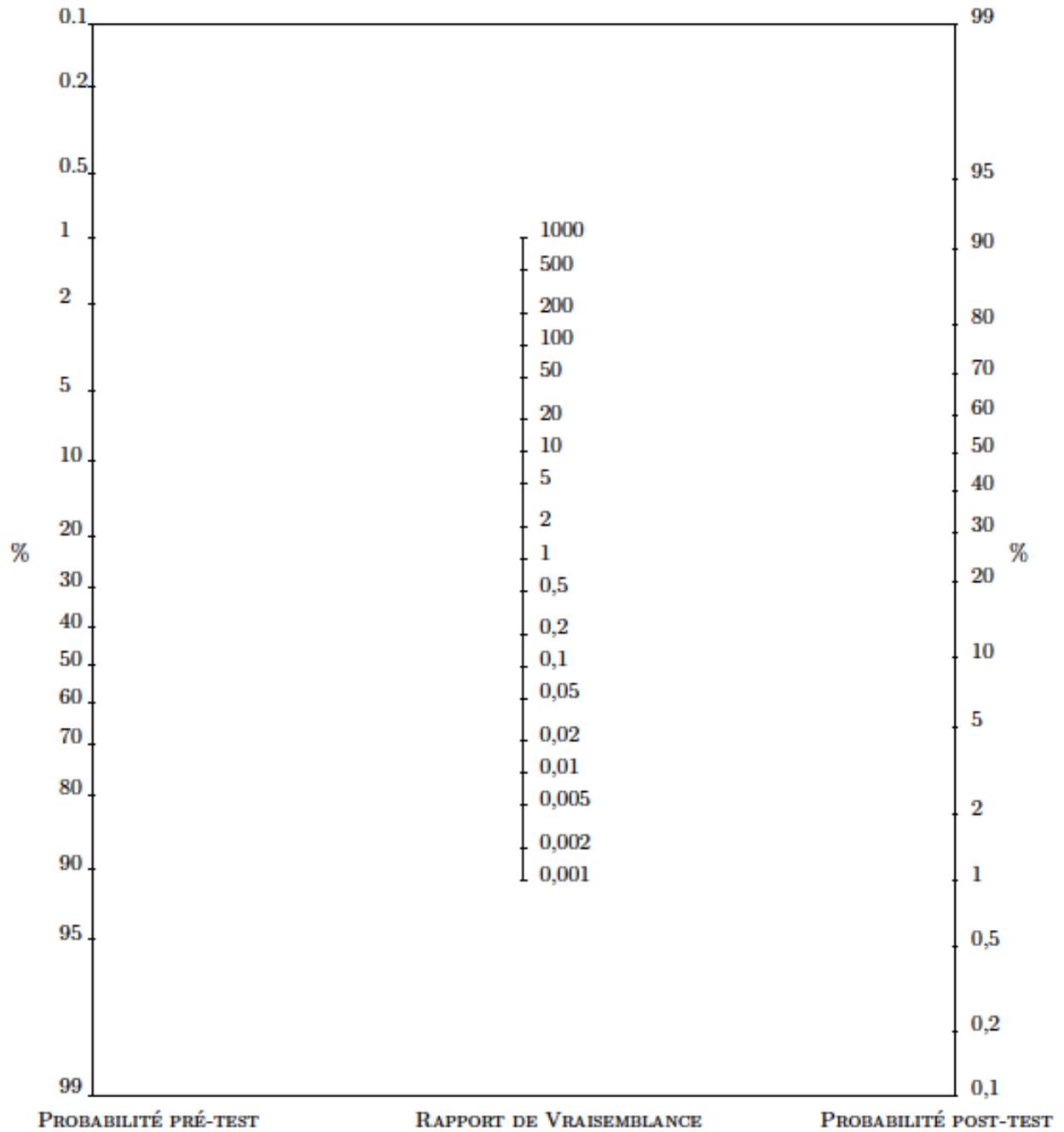
$$L = Se/(1-Sp) = (VP/malades)/(FP/non malades)$$

$$\lambda = (1-Se)/Sp = (FN/malades)/(VN/non malades)$$

Les rapports de vraisemblance sont indépendants de la prévalence de la maladie et sont un bon indice de la " valeur diagnostique " d'un test.

A partir de la probabilité pré-test d'être malade et du rapport de vraisemblance positive, le nomogramme de Bayes donne la probabilité post-test d'être malade. Pour cela, il faut prolonger sur l'axe de la probabilité post-test une droite passant par la probabilité pré-test et le rapport de vraisemblance positif.

Nomogramme de *Bayes*



Réalisé avec \LaTeX par N. Meyer.
D'après T.J. Fagan, NEJM;293:257.

GLOSSAIRE DE TERMES SOUVENT UTILISÉS DANS LA RÉDACTION DES ARTICLES MÉDICAUX

Nicolas MEYER^①, Jean-Louis IMBS^②

Ce glossaire a pour objectif de rappeler le sens des principaux termes utilisés dans la rédaction d'article médicaux. Il va à l'essentiel des notions nécessaires dans la méthodologie biostatistique tout en limitant l'utilisation des formules et cite les principales composantes de l'essai clinique contrôlé. Il a été complété par certaines définitions figurant dans le glossaire proposé sur le site de l'ECN.

Ajustement

Moyen, dans une enquête épidémiologique, de prendre en compte un facteur de confusion au moment de l'analyse.

Aléatoire

Dont la survenue dépend du hasard. L'attribution aléatoire d'un traitement fait confiance au hasard. Le tirage au sort des groupes de sujets permet de les rendre comparables pour tous les facteurs connus ou inconnus qui pourraient influencer sur le critère que l'on mesure. Souvent, cette hypothèse de comparabilité des groupes est vérifiée numériquement lors de l'analyse des résultats.

Analyse en intention de traiter

En toute rigueur, ce que l'essai contrôlé compare, c'est l'intention de traiter par un traitement A à l'intention de traiter avec un traitement B. Il faut donc conserver dans l'analyse tous les patients inclus dans l'étude selon leur répartition au hasard (randomisation). Cette allocation aléatoire est l'une des garanties que les deux groupes de patients (traité par A ou par B) sont comparables. Toute soustraction de l'analyse de patients (sortis d'étude, ou ne satisfaisant pas aux critères d'inclusion ou pour lesquels des valeurs sont manquantes, ou encore perdus de vue) risque d'induire un biais.

Cette approche prend la réalité clinique en compte : il est par exemple impossible de continuer à traiter un patient par un traitement qu'il ne supporte pas uniquement pour des exigences d'analyse statistique. En corollaire, il faut signaler que cette analyse ne permet que de mesurer l'efficacité clinique, c'est à dire une efficacité définie dans le cadre de la pratique courante. L'analyse en intention de traiter peut être associée avec une analyse selon le protocole (ou "per protocole") dans laquelle des patients peuvent être exclus de l'analyse pour des raisons variées comme celles citées ci-dessus.

Analyse per protocole

Cette analyse ne porte que sur les patients qui ont terminé et réalisé entièrement l'étude et pour lesquels toutes les mesures prévues sont disponibles.

^① Laboratoire de Biostatistiques et Informatique médicale – Faculté de Médecine – 11, rue Humann – 67085 Strasbourg Cedex
^② Institut de Pharmacologie – Faculté de Médecine – 11, rue Humann – 67085 Strasbourg Cedex

Analyse intermédiaire

Analyse effectuée avant l'inclusion de tous les sujets prévus. Elle est réalisée le plus souvent lorsque l'étude est longue. Elle doit être prévue dans le protocole, et le nombre de sujets nécessaires prend en compte le nombre d'analyses intermédiaires qui sont prévues. Pour chaque analyse intermédiaire, un seuil de signification doit être choisi.

Analyse séquentielle

L'analyse statistique d'une étude est classiquement réalisée une seule fois, à la fin de l'étude, lorsque toutes les données ont été recueillies. Dans la méthode séquentielle, l'analyse est réalisée après l'inclusion de chaque sujet afin de détecter le plus rapidement possible un produit efficace ou un produit dangereux ou inutile. Les études avec analyses intermédiaires rentrent dans le cadre des analyses séquentielles.

Ce type d'analyse nécessite des techniques statistiques particulières et permet en général de réduire le nombre de sujets nécessaire.

Appariement

Technique permettant de rendre comparables deux ou plusieurs groupes, en particulier par rapport à certains facteurs de confusion déjà connus dont on veut neutraliser les effets (voir appariement).

Appariement (former des paires)

Rendre comparables deux groupes en terme de facteurs de confusion potentiels. Pour chaque cas (exemple : un malade), on associe un ou plusieurs groupes témoins qui lui sont similaires pour un ou plusieurs facteurs (exemple : âge, sexe, niveau socio-économique).

Biais

Erreur systématique introduite par une erreur méthodologique. En présence d'un biais, l'échantillon sur lequel porte l'analyse n'est pas représentatif de la population à laquelle on s'intéresse.

Biais d'attrition

Le biais d'attrition intervient lorsque des patients randomisés sont exclus de l'étude ou écartés de l'analyse ou encore lorsque les sujets quittent d'eux même l'étude. Ce biais est fréquent et assez facile à détecter lorsque les effectifs sont suffisamment bien définis dans l'article. Dans la pratique ce biais est limité par une bonne organisation de l'étude.

Biais d'évaluation

Ce biais survient lorsque l'évaluation du résultats du traitement est influencée par des facteurs subjectifs l'étude n'étant pas en double insu (double aveugle). Il faut parfois faire intervenir un évaluateur externe pour mesurer objectivement le critère de jugement (encore appelé triple aveugle).

Biais de classement

Biais dans la mesure du facteur de risque ou dans la certitude de la maladie. Ce biais consiste à classer le facteur de risque ou la maladie dans une catégorie erronée. Cette erreur est quasi inévitable puisqu'aucun outil de mesure (interrogatoire, examen, test) n'est parfait.

Exemple : un comportement à risque minimisé par le malade, ou simplement non recherché dans le questionnaire.

Biais de confusion

Le biais de confusion est le biais entraîné par l'absence de la prise en compte d'un facteur pouvant influencer le résultat, de ce fait ce facteur tout en étant correctement réparti peut entraîner une dispersion importante cachant une différence réelle entre groupes. Des méthodes d'ajustement peuvent prendre en compte un facteur de confusion, encore faut-il l'avoir identifié. Un exemple classique en est le biais lié à l'exposition à l'alcool dans une étude sur l'effet du tabagisme dans les cancers ORL.

Biais de sélection

Ce biais existe lorsque les deux (ou plusieurs) groupe(s) construits ne sont pas comparables. C'est le cas lorsque l'on effectue une comparaison avec un groupe de référence (témoin) non construit à partir de la même population que celle des autres groupes. Deux points sont à souligner :

- La randomisation doit être effectuée selon les bonnes règles (p.ex. en centralisée si l'étude est multicentrique, équilibrée ou stratifiée si nécessaire).
- La bonne comparabilité doit être vérifiée sur certains facteurs car la randomisation ne donne pas une garantie absolue de comparabilité.

Biais de suivi

Ce biais est introduit lorsque la comparabilité n'est pas maintenue tout au long de l'étude. Les causes de ce biais sont nombreuses comme par exemple les déviations au protocole, les traitements concomitants qui peuvent être différents, les arrêts du traitement pour diverses raisons.

Caractérisation d'une grandeur qualitative

Une grandeur qualitative peut être à deux niveaux, comme la présence ou absence d'un signe (variable dichotomique), ou à plus de deux niveaux, comme l'évaluation de la gravité d'une maladie (variable polychotomique). La distribution d'une variable dichotomique suit une loi binomiale, celle d'une variable polychotomique suit une loi multinomiale. On peut caractériser de telles variables par la fréquence et l'écart type de chaque modalité. On peut également les caractériser par la fréquence et par l'intervalle de confiance de la fréquence.

Caractérisation d'une grandeur quantitative

Une grandeur quantitative est caractérisée par sa distribution (gaussienne ou autre), sa valeur moyenne et son écart type. D'autres paramètres sont parfois utilisés comme les valeurs extrêmes et les quantiles (quartiles, déciles). Une distribution différente d'une distribution gaussienne peut parfois être

transformée pour se ramener à une loi de Gauss et bénéficier des tests paramétriques. Lorsque aucune transformation n'est souhaitée, il convient d'utiliser des tests non paramétriques.

Causalité

Cette notion est employée lorsque l'on souhaite prouver que la variation d'un facteur explicatif est la cause des variations d'un facteur à expliquer. En recherche clinique, la seule méthode permettant de prouver une relation de cause à effet est la méthode expérimentale. Rechercher une relation de cause à effet entre deux variables nécessite la construction d'une étude expérimentale dans laquelle l'une des variables est contrôlée, la variable à expliquer étant simplement mesurée.

Remarque : L'existence d'une liaison entre deux variables obtenues autrement que par la méthode expérimentale ne prouve pas la causalité. Lorsqu'un coefficient de corrélation entre deux grandeurs quantitatives est statistiquement différent de zéro, on a uniquement montré que les deux variables n'étaient pas indépendantes (lorsque les distributions sont gaussiennes).

Cohorte de sujets

Groupe de personnes suivies dans le temps de manière prospective, ayant en commun d'avoir subi un événement semblable.

CPP

Comité de Protection des Personnes. Il est chargé de valider la conformité d'un protocole avec la législation, de vérifier le respect des principes éthiques et de vérifier que la sécurité des personnes se prêtant à la recherche est assurée notamment par l'évaluation de la balance bénéfice-risque. Tous les protocoles de recherche clinique doivent lui être soumis pour avis. Les études épidémiologiques sans investigations invasives sortent de ce champ. Il remplace le CCPPRB (Comité Consultatif de Protection des Personnes se prêtant à la Recherche Biomédicale) depuis 2006.

Comparabilité des groupes

Un des critères dans un essai contrôlé est la comparabilité des groupes. Les groupes doivent être obtenus par randomisation à partir d'une population initiale clairement définie. Il faut veiller à maintenir cette comparabilité afin d'éviter d'introduire des variations non contrôlées dans l'étude ce qui conduirait à des biais de suivi, d'attrition ou d'évaluation. Il est important de vérifier, une fois l'étude terminée et ses résultats saisis, qu'il n'y avait effectivement pas de différence entre les groupes au début de l'étude. Exceptionnellement, le hasard (la randomisation) peut faire qu'une telle différence existe, rendant alors très difficile l'interprétation de l'étude.

Comparaison de plusieurs groupes

Dans un essai contrôlé on peut être amené à comparer la moyenne de plusieurs groupes. La méthode à utiliser consiste à construire une étude basée sur l'analyse de la variance à un facteur à plusieurs modalités (plusieurs groupes). L'analyse de la variance a pour but de comparer les moyennes tout en gardant le risque α à 5%. La conclusion de l'analyse est que l'une ou plusieurs moyennes sont statistiquement différentes. L'analyse est complétée par la localisation de la différence (test de Newman-

Keuls, test de Scheffe). La mauvaise méthode consisterait à comparer les groupes entre eux (comparaison multiple).

Comparaison multiple

En répétant les tests de comparaison sur les mêmes données (chaque groupe est utilisé plusieurs fois pour le test) on modifie le risque α initial fixé. Par exemple, si le risque α est fixé à 0,05 pour un test, pour deux tests, le risque de conclure à tort à une différence sur au moins des deux tests sera de 0,08 et pour trois tests il sera de 0,11. Une façon de corriger cette modification du risque est de protéger le test selon la méthode de Bonferroni en baissant le risque α initial en fonction du nombre de tests prévus (prendre un risque $\alpha' = \alpha / n$ pour n tests).

Consentement éclairé

Libre acceptation (exprimée par écrit) d'une personne envisageant de participer à un essai donné. Cette acceptation formelle ne doit être demandée qu'après information (dite "éclairée") de la personne, précisant les objectifs, les bénéfices, les risques et les inconvénients potentiels, liés à l'essai ; la personne doit également être informée de ses droits et responsabilités. Son acceptation doit être recueillie avant son inclusion dans l'étude.

Critère de jugement (voir aussi critères d'efficacité)

L'essai contrôlé a pour objectif de décider si deux ou plusieurs groupes sont statistiquement différents pour une variable définie. Cette variable constitue le critère principal de jugement. Celui-ci peut être qualitatif ou quantitatif. Dans un essai contrôlé, il n'y a qu'un seul critère de jugement principal : c'est par rapport à lui que les effectifs nécessaires sont définis. D'autres paramètres que cette variable principale peuvent être intéressants à étudier : ils constituent des variables secondaires et aucun pari n'est fait sur elles.

Critères d'efficacité (en anglais : end points)

Le ou les critères permettant de mesurer les effets d'un médicament dépendent des buts de l'essai et, évidemment, de la classe du produit étudié. Un critère de jugement principal de l'efficacité doit être choisi de façon à répondre à l'objectif principal de l'étude. En somme, il s'agit du mode de mesure du bénéfice que le patient peut retirer du médicament.

Il existe trois types de critères d'efficacité : clinique, intermédiaire, de substitution.

Critères cliniques

Ils traduisent une amélioration quantitative ou qualitative de sa santé perçue par le patient, tels que :

- amélioration de la probabilité de survie ;
- prévention d'un état morbide ;
- disparition d'un symptôme douloureux, gênant ou invalidant ;
- guérison plus rapide ;
- amélioration de la qualité de la vie.

Critère d'inclusion / d'éligibilité / d'exclusion

Ensemble des critères spécifiant les caractéristiques des personnes à inclure ou à exclure de l'étude.

Critère intermédiaire

Amélioration du résultat d'un examen complémentaire paraclinique (radiologique par ex.) ou biologique (glycémie par ex.)

Critère de substitution (en anglais : surrogate end point)

Certains critères intermédiaires peuvent se substituer au critère clinique si :

- la preuve est faite que l'amélioration du critère de substitution coïncide certainement avec l'amélioration du critère clinique ;
- le critère de substitution est d'acquisition plus facile que celle du critère clinique.

Un exemple en est la baisse de la pression artérielle dans le traitement de la maladie hypertensive).

Critère composite

Le critère de jugement utilisé dans une étude est le plus souvent simple, c'est-à-dire qu'il n'est constitué que d'une seule mesure (un dosage biologique, une mesure échographique) ou événement (décès). On utilise parfois un critère composite, c'est-à-dire un critère qui est constitué de plusieurs éléments de jugement. Cela peut-être par exemple une survie sans événement, où on compte comme événement la rechute d'un cancer, sa progression, le décès ou l'apparition d'une métastase. Le critère composite peut-être obtenu par combinaison de plusieurs paramètres biologiques comme par exemple différents éléments d'un bilan lipidique.

Les critères composites ont l'avantage de donner une vision globale de la maladie et permettent parfois d'augmenter la puissance de l'étude mais ils sont souvent difficiles à interpréter, notamment lorsque les différents éléments qui le composent ont dans l'étude des évolutions discordantes.

Cross-Over

Essai thérapeutique où le sujet est pris comme son propre témoin. Un groupe de patients reçoit le traitement A puis le traitement B, l'autre groupe de patients reçoit le traitement B puis le traitement A.

Description d'une population

La description de la population étudiée permet au lecteur d'apprécier le contexte dans lequel l'étude est effectuée. Elle est nécessaire afin de connaître avec précision les critères d'inclusion et d'exclusion de l'étude, le lecteur pourra ainsi évaluer la portée de l'étude. Une description complète facilitera la comparaison des résultats avec d'autres études et pourra faciliter une méta analyse.

La population étant définie par le problème étudié, l'étude sera en général effectuée sur un échantillon représentatif de cette population.

Différence de risque : voir Mesures de risque

Degré de signification (ou de significativité)

Risque (au sens de pari) de se tromper lorsque l'on conclut à une différence dans un test statistique. Lors de la réalisation d'un test statistique pour comparer deux groupes, on peut par erreur conclure à une différence qui n'existe pas dans la réalité. Lors du test, on calcule la probabilité d'observer une différence au moins aussi importante que celle que l'on a obtenue : c'est le degré de signification " p ". Si cette probabilité est inférieure à un seuil classiquement fixé à $\alpha=5\%$ (0,05), on conclue que les deux groupes diffèrent. En procédant de cette manière, on se trompera dans 5% des cas en déclarant une différence alors qu'il n'y en a pas.

Double – insu (ou double aveugle, traduit littéralement de l'anglais "double-blind")

Cette notion s'applique à l'essai contrôlé organisé de manière que ni le patient, ni le médecin, ni le personnel soignant, ni l'entourage du malade ne connaissent celui des traitements testés que reçoit un patient donné. En revanche, les médicaments ou produits testés dans l'essai considéré sont bien sûr connus de tous, y compris des patients qui ont été "éclairés" par une information appropriée. L'essai contrôlé en double insu a pour objet de limiter les biais qui pourraient s'introduire dans l'étude par la connaissance du traitement car patients et médecins pourraient alors manquer d'objectivité.

Ecart-type

Cette notion caractérise la variabilité d'une grandeur quantitative, c'est-à-dire la dispersion d'une série de mesure quantitative. Cette grandeur est utilisée dans de nombreux calculs en statistique, par exemple pour calculer un intervalle de confiance, définir un paramètre discriminant pour effectuer un test, calculer des effectifs nécessaires, etc...

Echantillon

L'objectif de la recherche clinique est d'obtenir des résultats (des connaissances) présentant un caractère général (universel) afin de pouvoir utiliser cette connaissance sur l'ensemble d'une population. L'étude qui amènera cette connaissance ne peut que très rarement être effectuée sur toute la population, c'est pour cette raison qu'elle sera effectuée sur un échantillon représentatif de la population. Constituer un échantillon, c'est tirer au hasard une sous-population d'effectif n dans une population (population parente) de taille N .

La précision des paramètres estimés augmente avec la taille de l'échantillon.

C'est à partir des connaissances obtenues sur l'échantillon que l'on induit les connaissances sur la population. On détermine certaines grandeurs à partir de l'échantillon, celles-ci constituent des estimations des grandeurs correspondantes de la population. Lorsque l'échantillon est bien constitué, la statistique inductive nous donne des indications sur la population avec un certain risque d'erreur. On peut ainsi, à partir de l'échantillon, caractériser la population en effectuant une estimation ponctuelle de certaines grandeurs ou une estimation par intervalle de confiance de ces mêmes grandeurs. Cette démarche nécessite au préalable la vérification de certaines hypothèses.

Remarque : Tirer au hasard ou au sort, ou randomiser (de " random ", hasard en anglais), permet d'éliminer toute subjectivité (biais) dans le choix de l'échantillon. Dans la pratique, pour obtenir un échantillon correct d'une population définie par des critères d'inclusion et d'exclusion, on peut se servir de tables numériques fournissant des séries de nombres aléatoires ou utiliser des programmes informatiques.

Effectifs nécessaires : voir " nombre de sujets nécessaire "

Effet *nocebo*

Effet négatif de la prise d'un médicament qui n'est pas lié aux protéines physico-chimiques de la molécule, mais au fait même que l'on prend un traitement (effet psychologique).

Effet *placebo*

Effet positifs de la prise d'un médicament qui n'est pas lié aux propriétés physico-chimiques de la molécule, mais au fait même que l'on prend un traitement (effet psychologique).

Enquête longitudinale

Enquête au cours de laquelle des informations sont recueillies de façon longitudinale, c'est-à-dire tout au long de l'étude. La durée de l'étude est définie et peut être assez longue (plusieurs années). Elle peut être prospective ou rétrospective.

Enquête transversale

Enquête qui consiste à recueillir simultanément ou quasi simultanément les données relatives à la maladie et aux facteurs de risques étudiés.

Essai contrôlé (Essai comparatif randomisé, Essai thérapeutique)

Type d'étude demandant une organisation bien codifiée dans laquelle les échantillons retenues pour l'étude proviennent d'une population commune par randomisation. Certaines grandeurs sont contrôlées par la définition de la population, les autres grandeurs étant réparties de façon aléatoire dans les groupes. Dans un tel essai, on peut mettre en évidence des relations de cause à effet. Cette étude est démonstrative, la différence mise en évidence provient du facteur étudié.

Selon l'organisation, l'essai peut être en simple aveugle (simple insu) ou en double aveugle (double insu) selon que le patient ou le patient et le médecin ignorent quel est le médicament attribué.

Essai de non infériorité¹

Il existe de plus en plus de traitements reconnus, efficaces, validés. Un nouveau médicament dans une classe thérapeutique où existe un tel traitement de référence ne pourra pas, éthiquement, être comparé à un placebo dans le cadre d'un essai de supériorité. En aucun cas, le non rejet de l'hypothèse H_0 (postulant que les traitements sont identiques) dans un essai de supériorité ne permet de conclure à une

¹ Note : Les essais de non infériorité font intervenir un domaine d'équivalence unilatéral défini par ce seuil. C'est un domaine d'équivalence bilatéral qui est fixé dans les " essais d'équivalence ", non traités ici car rarement utilisés dans les essais thérapeutiques

égalité. C'est pour répondre à cette situation à présent fréquente que la méthode des essais de non infériorité s'est développée.

Actuellement, les avancées thérapeutiques portent souvent sur une facilité ou un confort d'utilisation, une meilleure sécurité d'emploi ou un moindre coût, alors que l'efficacité est similaire au(x) médicament(s) de référence. Ces avancées peuvent être jugées suffisantes pour permettre d'accepter une certaine perte d'efficacité. Il s'agira alors de montrer que la différence d'efficacité entre les traitements comparés (nouveau médicament vs médicament de référence) est inférieure à un seuil préalablement fixé. Ce seuil correspond à la quantité d'efficacité qu'il est jugé acceptable de perdre face aux autres avantages du nouveau médicament. C'est le plus souvent une limite unilatérale qui sera fixée par ce seuil définissant la non infériorité. Cette méthodologie est très particulière à cette situation, clairement distincte des essais de supériorité.

Le choix du seuil de non infériorité relève d'une discussion d'experts pesant les avantages potentiels du nouveau médicament face à la perte d'efficacité acceptée. Il dépend de la classe thérapeutique, du critère de jugement principal, de l'efficacité du médicament de référence telle qu'elle est reconnue au moment de l'étude, et doit être défini à l'avance. En général, on conserve au moins 75 % de l'effet du traitement de référence par rapport à celui du placebo.

Plusieurs exigences méthodologiques découlent logiquement de ces concepts :

- Les résultats sont exprimés par l'intervalle de confiance à 95 % de la moyenne de la différence d'efficacité sur le critère de jugement principal entre les deux traitements. La non infériorité est démontrée si cet IC_{95%} ne dépasse pas le seuil de non infériorité.
- Il est nécessaire d'assurer une puissance maximale à l'étude de façon à éviter tout biais qui amoindrirait une différence sinon significative (choix des posologies, limitative du nombre de perdus de vue, homogénéité des sorties d'étude entre les bras, calcul du nombre de sujets). Rappelons que la façon la plus sûre de ne pas montrer de différence entre deux traitements est un essai de supériorité de mauvaise qualité.
- L'efficacité du médicament de référence doit correspondre aux données de la littérature (validation externe).

Essai en double aveugle

Essai au cours duquel, ni le patient, ni le médecin ne connaissent le traitement pris. Cela permet d'éliminer l'effet placebo chez le patient et les biais de mesure liés à la subjectivité du médecin. Dans l'essai en triple aveugle, le chercheur qui analyse les résultats ne sait pas quel groupe de patients a reçu quel type de traitement.

Essai en simple aveugle

Essai au cours duquel le patient ne connaît pas le traitement qu'il reçoit. Le médecin connaît le traitement que chaque patient reçoit. Cela permet normalement de neutraliser les effets placebo et *nocebo*.

Essai ouvert

Essai thérapeutique souvent mené sur petit groupe de sujets, parfois comparatif, permettant d'étudier la faisabilité d'un essai comparatif à la plus grande échelle.

Etude cas – témoins ; Enquête cas - témoins

Une telle étude est classée dans les enquêtes. Une enquête n'est jamais démonstrative.

Dans une enquête de ce type, on connaît les cas (les malades sont souvent ceux de nos hôpitaux) et l'on cherche des témoins (non malades) ressemblants le plus possible aux malades. On dispose alors de deux groupes, plus ou moins comparables, sur lesquels on recherche le facteur à étudier, par exemple l'exposition à un risque. Ce type d'enquête permet de construire une table à quatre cellules contenant les effectifs pour les situations M+E+, M+E-, M-E+, M-E-. On peut calculer s'il existe une liaison entre le facteur maladie (M) et le facteur exposition (E). Etant donnée la structure de l'ECT, on ne peut pas calculer le risque relatif. Par contre, on peut calculer le rapport des cotes ou Odds Ratio (voir ce terme dans les mesures de risque) qui converge vers le risque relatif lorsque la maladie est rare. Ce type d'enquête est plus facile à réaliser que l'enquête sur échantillon représentatif. Il est l'objet cependant d'un certain nombre de biais.

Ce type d'étude est fréquemment effectué chez les patients hospitaliers, certains patients malades étant pris comme " cas ", d'autres patients non malades étant pris comme " témoins ". Cette situation pose souvent de nombreux problèmes d'interprétation.

	M +	\bar{M} -
E +	a	b
\bar{E} -	c	d

Les termes a, b, c, d représentent des effectifs. L'effectif total est (a + b + c + d).

Il faut remarquer que les taux de malades et de non malades ne sont pas représentatifs de la population par conséquent, les rapports [a / (a + b)] et [c / (c + d)] n'ont pas de sens (double trait vertical rappelant cette notion, donc pas de calcul du risque relatif mais de l'odds ratio).

Etude d'une série de cas

Ce type d'étude correspond à l'analyse d'une série de cas sans pour autant avoir un objectif de comparaison avec un groupe témoin. L'analyse d'une telle série relève essentiellement de la statistique descriptive. Elle permet de faire le point sur une méthode dans un but descriptif ou dans un objectif préliminaire de faire une étude plus structurée dans un deuxième temps. On accorde généralement un faible niveau de preuve à ce type d'étude (4 sur une échelle allant de 1 à 4 (critères de l'ANAES)).

Etude de cohorte ; Enquête de cohorte ; étude longitudinale

Une telle étude est classée dans les enquêtes. Une enquête n'est jamais démonstrative.

Ce type d'enquête encore appelé "enquête exposés - non exposés" dans laquelle on suit des individus dans le temps, les uns étant exposés à un risque et les autres non. On observe l'arrivée des maladies et au bout d'un certain temps d'observation, on effectue l'analyse. Ce type d'enquête permet de construire une table à

quatre cellules contenant les effectifs pour les situations M+E+, M+E-, M-E+, M-E-. Il est possible de déterminer s'il existe une liaison statistique entre les deux facteurs (calcul du χ^2 : Chi²) et l'on peut calculer le risque relatif. La liaison obtenue n'est cependant pas une liaison de cause à effet. Ce type d'enquête est plus facile à réaliser que l'enquête sur échantillon représentatif, il demande cependant une longue période d'observation pendant laquelle la comparabilité initiale s'estompe fréquemment.

	M +	\bar{M} -
E +	a	b
\bar{E} -	c	d

Les termes a, b, c, d représentent des effectifs. L'effectif total est (a + b + c + d).

Etude sur échantillon représentatif ; Enquête sur échantillon représentatif

Une telle étude est classée dans les enquêtes. Une enquête n'est jamais démonstrative.

Après avoir défini la population d'étude, le tirage au sort sera effectué. On construira une table selon le schéma ci-dessous selon les facteurs maladie et exposition (facteur E supposé être lié à la maladie M). Ce type d'enquête permet de construire une table à quatre cellules contenant les effectifs pour les situations M+E+, \bar{M} +E-, M-E+, \bar{M} -E-. Il est alors possible de déterminer s'il existe une liaison statistique entre les deux facteurs (calcul du Chi²) et calculer le risque relatif. La liaison obtenue n'est cependant pas une liaison de cause à effet.

	E +	\bar{E} -
M +	a	b
\bar{M} -	c	d

Les termes a, b, c, d représentent des effectifs. L'effectif total est (a + b + c + d).

Ce type d'enquête est souvent difficile à mettre en place et demande de grands effectifs lorsque la maladie est rare dans la population.

Etude pronostique

Formuler un pronostic, c'est estimer les probabilités des divers modes d'évolution d'une maladie. L'étude pronostique doit s'attacher à analyser l'évolution d'un groupe de sujet à partir d'une date ou d'un événement. Les études pronostiques s'appuient sur les études de cohorte et de survie.

Etude transversale ; Enquête transversale

C'est une étude dans laquelle l'exposition au facteur de risque et la maladie sont mesurées en même temps. Ce type d'étude permet de donner une vision à un instant donné. Elle permet d'obtenir la prévalence de diverses maladies.

Elle a un certains nombre d'inconvénients, en particulier elle ne permet pas d'apprécier la séquence " Exposition au risque → maladie "

Facteur de risque

Un facteur de risque est un facteur dont la présence est associée statistiquement à une augmentation de la fréquence d'une maladie. Lorsque le lien entre le facteur et la maladie n'est pas causal, on dit que le facteur est un marqueur de risque pour préciser que le facteur, si sa présence indique un plus grand risque d'avoir la pathologie en question, n'est pas forcément la cause de cette pathologie.

Hazard ratio : voir Mesures de risque

Incidence (taux d')

Fréquence des cas nouveaux dans une période de temps donnée.

Intervalle de confiance

Cette notion s'adresse aux grandeurs estimées à partir d'un échantillon. Nous considérons ici uniquement la fréquence d'un caractère ou la valeur moyenne d'une grandeur. La statistique inductive nous donne des informations sur la population à partir des informations obtenues sur l'échantillon. L'estimation d'une grandeur à partir d'un échantillon est entachée d'une incertitude puisque d'autres échantillons issus de la même population donneraient d'autres estimations. Lorsqu'un certain nombre de conditions sont réunies (par exemple population gaussienne ou effectif supérieur à 30) on peut trouver un intervalle tel qu'il y a 95% de chances pour que la vraie valeur du paramètre de la population (fréquence du caractère dans la population ou moyenne de la grandeur étudiée dans la population) soit dans l'intervalle. Cet intervalle est appelé "intervalle de confiance".

On peut et il faut calculer l'intervalle de confiance pour tout paramètre statistique et pour toute mesure de risque. Ces intervalles de confiance doivent normalement être présents dans un article correctement rédigé.

Exemple : OR = 1.53 [1.08 – 2.2]. Les valeurs entre les crochets indiquent respectivement la borne basse et la borne haute de l'intervalle de confiance de l'odds-ratio obtenu dans l'étude. Ces valeurs traduisent que la l'étude aurait pu générer avec 95% de chance des valeurs d'OR allant de 1.08 à 2.2.

Indépendance

Neutralité d'un événement A sur un événement B : deux événements sont indépendants si l'issue de l'un n'influe pas sur l'issue de l'autre.

Intervention

Une intervention est l'application d'une mesure (au sens large) destinée à améliorer la santé d'une population. La mise en place d'une campagne de vaccination, l'utilisation d'une nouvelle thérapeutique, la modification de l'accès aux soins sont des exemples d'intervention.

Latéralité d'un test statistique

Un test statistique est soit bilatérale, lorsque l'on ne sait pas a priori quelle est le sens de la différence entre les groupes que l'on observera, soit unilatérale lorsque la différence ne peut se faire, pour certaines raisons, que dans un sens ou n'a d'intérêt que dans un sens.

Médicament

Selon l'OMS (Organisation Mondiale de la Santé), est un médicament toute substance ou composition administrée pour guérir ou prévenir une maladie ou un symptôme, établir un diagnostic, modifier ou restaurer les fonctions de l'organisme.

Méthode expérimentale

L'expérience scientifique est une comparaison où tous les facteurs confondants (par ex. pour l'essai clinique : régression à la moyenne, effet placebo, évolution spontanée, terrain pathologique particulier, précision des instruments de mesure des critères de jugement), connus ou inconnus, jouent de la même manière sur les termes de la comparaison. L'hypothèse testée sur laquelle s'appliqueront les tests statistiques, doit avoir été parfaitement formulée avant le début de l'expérience.

L'essai contrôlé est l'application à la recherche clinique des principes de la méthode décrite par Claude Bernard.

Mesures de risque

Considérons deux groupes de sujets : l'un est exposé à un facteur et l'autre n'est pas exposé. On souhaite savoir si l'exposition au facteur augmente la fréquence de la maladie par rapport à l'absence d'exposition. Pour comparer les deux groupes, on calcule différents indices, appelés de façon générique " mesures de risque ". Ces mesures de risques sont :

Le Risque Absolu (RA)

Le Risque Relatif (RR)

La Différence de Risque = Risque attribuable = Réduction absolue du risque (DR)

La Réduction Relative du Risque (RRR)

Le nombre de sujets à traiter (NST) (en anglais " *Number Needed to Treat, NNT* ")

L'Odds-Ratio (OR)

Le rapport de risque (Hazard Ratio, HR)

Les cinq premières mesures sont utilisées dans les études prospectives tandis que la sixième peut-être utilisée dans les études prospectives ou rétrospectives (études cas-témoins).

Ces différentes mesures peuvent être résumées à partir du tableau ci-dessous. Les deux lignes correspondent aux deux niveaux d'exposition (exposition présente / absente) que l'on peut prendre au sens large et donc remplacer par l'attribution du traitement à l'étude ou d'un traitement de référence. Les deux colonnes correspondent au résultat de l'étude : sujets sains ou malades, que l'on peut remplacer de façon plus générale par la présence ou l'absence d'un événement d'intérêt.

	M+	M-
E+	a	b
E-	c	d

Les termes " a, b, c, d " représentent des effectifs. L'effectif total est (a + b + c + d). Les termes a / (a + b) et c / (c + d) représentent les risques absolus (taux d'incidence) d'être atteint par la maladie dans l'un et l'autre groupe.

Les mesures de risque se calculent de la façon suivante :

Le **risque absolu** : c'est le risque de développer/de présenter la maladie dans un groupe d'exposition donné. Le risque absolu chez les exposés est donc : $RA^{E+} = a/(a+b)$ et chez les non-exposés, le risque absolu vaut : $RA^{E-} = c/(c+d)$. Le risque absolu n'est donc pas comparatif, par définition, contrairement aux autres mesures de risques. Ces deux risques absolus sont la base du calcul de plusieurs mesures de risques : RR, DR, RRR et NST.

Le **risque relatif**, noté RR :

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

Le risque relatif peut être calculé sur une étude de cohorte ou une étude sur échantillon représentatif.

La **différence de risque**, DR :

$$DR = [a/(a+b)] - [c/(c+d)]$$

C'est la fréquence chez les exposés – la fréquence chez les non exposés. Selon les besoins, on peut bien sûr calculer $DR = [c/(c+d)] - [a/(a+b)]$ pour avoir une différence de risque de sens opposée (positive au lieu d'être négative ou vice-versa).

La DR est aussi appelée **risque attribuable**, RA lorsque l'on peut donner un rôle causal au facteur de risque.

La **réduction relative du risque**, RRR :

$$RRR = \frac{a/(a+b) - c/(c+d)}{c/(c+d)} = RR - 1$$

C'est la part de la différence de risque rapportée au risque dans le groupe non exposé. Il est en général utilisé dans les essais thérapeutiques où il mesure la diminution relative du risque en présence du traitement étudié.

Le **nombre de sujets à traiter**, NST (en anglais : " *Number needed to treat* "):

C'est le nombre de sujets à traiter pour éviter l'apparition d'un cas supplémentaire.

$$NST = \frac{1}{DR}$$

Remarque : dans cette formule, la différence de risque DR doit être exprimée par une valeur comprise entre 0 et 1 et pas par un pourcentage, qui varie de 0 à 100%. Donc, pour une différence de risque de 13%, le $NST = 1/0.13 = 7.8$, soit huit sujets. Il faut traiter en moyenne 8 sujets pour éviter un évènement.

L'**Odds-Ratio**, OR :

$$OR = \frac{ad}{bc}$$

On peut le calculer dans une étude prospective (étude de cohorte ou essai thérapeutique, même si ce n'est pas le meilleur indice) ou rétrospective (cas-témoins). Il constitue une bonne approximation du risque relatif quand la fréquence de la maladie est faible.

Tous ces termes sont des estimations obtenues à partir d'un échantillon d'individus, il convient dès lors de leur associer un **intervalle de confiance**.

Le **rapport de risque** (ou hazard ratio, HR) est un rapport de taux d'incidence. Il est obtenu lors de la comparaison de la durée de survie entre deux groupes. Il ne faut pas le confondre avec le risque relatif même s'il peut s'interpréter de manière similaire.

Mesures répétées

Dans certaines études, une même mesure est réalisée de manière itérative chez un même sujet.

L'analyse de ce type de donnée implique des méthodes statistiques adaptées à cette situation.

Méta-analyse

Une méta-analyse est une "analyse d'analyse". Elle consiste à colliger les données de plusieurs études réalisées sur un même sujet avec des protocoles les plus proches possibles afin de profiter de la puissance apportée par l'augmentation du nombre de cas. Elle tente de trancher des controverses apparues à la suite d'essais contradictoires. Elle tient compte de la variabilité des résultats des différentes études.

Méthodes multivariées

Les modèles étudiés en recherche clinique sont complexes, une grandeur dépend en général de plusieurs facteurs. On cherche à effectuer une analyse "globale" en expliquant une variable (variable à expliquer) en fonction de plusieurs autres variables (variables explicatives). De telles méthodes sont appelées multivariées, nous citons par exemple la régression multiple, la régression logistique, le modèle de Cox pour l'étude des facteurs pronostics dans la survie, etc...

Ces méthodes sont très accessibles aujourd'hui et doivent être utilisées lorsque cela s'avère utile. Elles doivent cependant faire suite aux méthodes univariées qui permettent de comprendre plus facilement les relations pouvant exister entre grandeurs.

Il faut remarquer que dans toute étude on cherche à approcher la réalité médicale et biologique à travers des modèles définis. Il faut s'assurer que le modèle utilisé est capable de décrire le phénomène et par conséquent en valider son utilisation. Ceci est une démarche qui reste difficile car de nombreux facteurs tels que le modèle lui-même (additif, multiplicatif) et le codage des variables peuvent intervenir.

Moyenne

Il s'agit de la moyenne arithmétique définie comme la somme des valeurs divisée par l'effectif. Cette grandeur est le paramètre majeur caractérisant un ensemble de mesure ou plus classiquement une distribution. Cette valeur a son importance puisque dans les modèles concernés par la recherche clinique on recherche une différence de valeur moyenne entre deux ou plus de deux groupes lorsque la variable est quantitative. Le test statistique est construit pour tester une différence entre les valeurs moyennes des groupes.

Nombre de sujets à traiter : voir Mesures de risque

Nombre de sujets nécessaire

L'objectif d'un essai thérapeutique est de conclure à une différence entre les effets mesurés dans les deux groupes au seuil de risque $\alpha = 0,05$ de conclure à tort que la différence existe. Ne pas conclure lorsque le test est non significatif peut signifier que la différence n'existe pas ou que le test manque de puissance. Pour éviter cette situation, il faut contrôler la puissance statistique du test par un calcul des effectifs nécessaires.

Le calcul du nombre de sujets nécessaire est basé sur les paramètres suivants :

- le risque α (de conclure à une différence qui n'existe pas) accepté (le plus souvent à 0,05, mais d'autres valeurs sont possibles) ;
- le risque β (de ne pas saisir une différence alors qu'elle existe) accepté (le plus souvent à 0,10 ou 0,20 (ce qui correspond à une puissance statistique de 90 ou 80%) ;
- la différence jugée intéressante sur le plan clinique entre l'effet thérapeutique observé dans chacun des deux groupes ;
- la variabilité du critère de jugement (d'efficacité) principal, autrement dit la dispersion des valeurs mesurées sur l'échantillon.

Remarques :

- L'importance de définir un objectif principal, impliquant l'utilisation d'un critère de jugement principal pour sa mesure, apparaît parfaitement : le calcul du nombre de sujets nécessaire sera fait selon la variabilité de cette mesure (et non de celle d'autres critères de jugement ou d'efficacité, dits alors secondaires) et la puissance de l'étude en découle directement.
- Il peut arriver que la dispersion des mesures du critère de jugement principal ne soit pas connue lors de la mise au point du protocole. Deux solutions sont alors possibles :
 - un calcul a posteriori vérifiant à partir des mesures effectuées que le nombre de sujets inclus dans l'étude permet d'obtenir une puissance statistique suffisante ;
 - la réalisation d'une pré-étude permettant d'estimer grossièrement cette variabilité.

- Il faut se méfier d'un excès d'optimisme dans la différence à mettre en évidence car une différence surestimée risque de ne pas permettre à terme de conclure par manque de puissance.
- Le calcul des effectifs nécessaires demande la connaissance de certaines grandeurs assez faciles à obtenir dans les cas de comparaisons entre deux groupes, mais beaucoup plus difficile à obtenir dans des plans d'étude complexes, comme par exemple dans l'analyse de la variance à plusieurs facteurs ou de nombreuses grandeurs sont à connaître.
- Une fois le calcul du nombre de sujets fait, il est habituel d'augmenter le effectifs de 10 ou 15% de façon à compenser les éventuelles sorties d'étude.

Objectif principal, secondaire

Toute étude est normalement assortie d'un objectif principal, ce pour quoi on réalise l'étude. Cet objectif est par exemple de connaître le taux de survie d'un certain groupe de patient atteint d'un cancer donné ou bien de comparer un taux de guérison pour deux traitements. La définition de l'objectif principal est importante car elle permet de spécifier la mesure qui servira de critère de jugement principal. Ainsi, pour connaître le taux de survie de patients atteints d'un cancer, il faudra définir et calculer un taux de survie à 1 an ou 5 ans par exemple et recueillir les éléments permettant de calculer ce taux.

L'objectif principal est unique (le plus souvent) et sert à déterminer le nombre de sujets nécessaire via le critère de jugement principal lui correspondant.

Des objectifs secondaires sont très fréquemment utilisés. Ils ne servent pas à calibrer l'étude mais permettent d'explorer certains aspects du problème.

Observance (en anglais : compliance)

Degré de coïncidence entre le comportement d'une personne et les conseils et prescriptions donnés par son médecin.

Odds ratio : voir Mesures de risque

Perdus de vue

Malades " disparus " avant la fin de l'essai et pour lesquels aucune information n'est plus disponible depuis cette " disparition ". Cette situation est différente de celle où le malade reste accessible mais pour lequel manquent certaines informations (valeurs manquantes). Un pourcentage de $\geq 5\%$ de perdus de vue dans un essai est inacceptable.

Petits effectifs

La notion de petits effectifs est associée à un calcul de χ^2 dans lequel on considère qu'il faut des effectifs théoriques au moins égaux à 5 pour être dans des conditions acceptables de calcul. Lorsque les effectifs théoriques sont plus petits que 5 on considère que l'on est dans le cas des " petits effectifs ". On utilise alors le calcul des probabilités exactes de Fisher pour pallier cette difficulté. Ce calcul est dit " exact " car il ne fait pas appel à des méthodes asymptotiques.

Il faut remarquer que si ces calculs sont tout à fait utilisables en recherche clinique, leur résultat est peu stable au regard de la taille de l'échantillon et les conclusions sont donc très fragiles.

Plan d'expérience (plan expérimental)

Les plus courants sont :

- **plan en groupes parallèles :**

Les participants sont randomisés en 2 (ou plus) groupes dont chacun va recevoir un des traitements testés.

- **plan en chassé-croisé (ou cross-over) :**

Chaque sujet est son propre témoin : il reçoit un traitement A pendant une première période et l'autre traitement B pendant une deuxième période, l'allocation aléatoire s'appliquant à l'ordre A-B ou B-A de sorte qu'autant de patients recevront d'abord B ou A.

- **plan factoriel :**

Utilisé pour évaluer la contribution de chacun des traitements prescrits en association. Par exemple, un plan factoriel conçu pour évaluer l'effet antihypertenseur l'association d'un bêta-bloquants et d'un diurétiques comportera 4 groupes de traitements (parallèles) en double insu : association bêta-bloquant + diurétique ; bêta-bloquant + placebo du diurétique ; diurétique + placebo du bêta-bloquant ; placebo du bêta-bloquant + placebo du diurétique.

Prévalence

Dans une population, la prévalence d'une maladie est la proportion de sujets présentant la maladie à un instant t .

Principe du test statistique

La comparaison de deux traitements est réalisée par une stratégie définie dans le principe du test statistique. En comparant deux échantillons, par exemple un échantillon traité par A et un échantillon traité par B, il s'agira de dire s'ils proviennent de la même population ou non (ou que les deux traitements ne montrent pas de différence ou non). Cette comparaison est effectuée à l'aide de mesures issues de chaque échantillon, une fréquence ou une moyenne par exemple.

Dans le principe du test statistique, on pose une hypothèse H_0 formulant que les deux échantillons proviennent de la même population. On calcule ensuite une grandeur discriminante à partir des différences constatées sur les grandeurs étudiées.

Lorsque la différence constatée est grande, on rejette l'hypothèse H_0 et l'on admet que les deux échantillons ne proviennent pas de la même population, par conséquent que A est différent de B. La statistique nous apprend que cette conclusion est effectuée sous le risque $\alpha = 0,05$ d'avoir rejeté H_0 à tort.

Lorsque cette différence est trop faible, on garde l'hypothèse H_0 et l'on admet que les deux échantillons proviennent de la même population : on dit qu'aucune différence significative entre A et B n'a été mise en évidence. La théorie statistique montre que cette conclusion est effectuée sous le risque β de ne pas avoir mis la différence en évidence si elle existe. Ce risque peut être diminué en augmentant les effectifs.

Puissance d'un essai clinique

C'est l'aptitude d'un essai à mettre en évidence l'effet du médicament testé. Cette puissance relève de quatre composantes :

- Puissance statistique :

Le nombre de malades inclus est un élément essentiel de la puissance d'un essai clinique. La puissance statistique ($1-\beta$) est la probabilité de trouver une différence qui existe. Elle est complémentaire du risque β ou risque de 2^e espèce qui est le risque de ne pas trouver une différence qui existe. Dans une étude, le calcul du nombre de sujets est réalisé de manière à assurer la puissance statistique souhaitée. Le nombre de sujets augmente avec la puissance et vice-versa.

- Puissance pharmacologique :

Elle est inhérente à la logistique de l'essai. Ainsi, il faut que le médicament à tester et médicament de référence soient administrés à posologies adéquates, que l'observance des traitements soit satisfaisante, que la fréquence des prises soit adaptée à la pharmacocinétique des médicaments testés.

- Puissance méthodologique :

Le plan d'expérience doit être adapté à la situation explorée : par ex. un essai en chassé-croisé ne peut être utilisé lorsque la pathologie traitée varie spontanément et rapidement. La qualité de la réalisation de l'étude intervient également : sorties d'études, valeurs manquantes, participants perdus de vue, sont autant de paramètres susceptibles de réduire la puissance d'une étude.

- Puissance analytique :

La méthode statistique doit être adaptée au protocole choisi et aux données recueillies. Ainsi, l'utilisation répétée des mêmes variables pour des comparaisons multiples demande une adaptation du seuil du risque α et entraîne une diminution de la puissance éventuellement compensée par une augmentation des effectifs de l'échantillon.

Randomisation

C'est une méthode pour diviser une population en groupes de telle manière que l'affectation des sujets laisse les groupes comparables. Chaque sujet a les mêmes chances d'être affecté à l'un ou l'autre des groupes. La randomisation permet d'éviter le biais d'attribution et par conséquent de conclure à l'action du traitement lorsqu'une différence significative apparaît. Lors d'une randomisation, les allocations d'individus doivent être définies à l'avance. On pourra se servir de tables de randomisation pour affecter les individus aux divers groupes (dans le cas de deux groupes l'un des groupes recevra le traitement A, l'autre le traitement B).

Dans certains cas on peut utiliser des méthodes permettant d'équilibrer les effectifs tous les n patients (randomisation par blocs) ou randomiser pour représenter de façon équilibrée un paramètre dans les groupes (par exemple l'âge) en randomisant par stratification.

Comparer des patients recevant un nouveau traitement à ceux recevant un ancien n'est pas une étude randomisée, il est presque certain que des biais sont présents dans l'étude.

Réduction absolue du risque : voir Mesures de risque

Réduction relative du risque : voir Mesures de risque

Référence (médicament ou traitement de)

C'est le traitement du groupe servant de contrôle. Son choix dépend de la situation thérapeutique étudiée :

- il n'existe aucune thérapeutique pour la maladie considérée : placebo ;
- il existe un traitement établi (soit par des essais contrôlés antérieurs, soit par une longue et indiscutable expérience clinique) : médicament validé à la posologie active, ou, si l'éthique le permet, un placebo.

Régression à la moyenne

Phénomène statistique obligatoire, dont l'intensité est liée à la variabilité intra-sujet de la variable considérée, observé chaque fois qu'est suivie l'évolution d'un paramètre au sein d'un échantillon tronqué (i.e. limité par un seuil) d'une population.

Ainsi, l'identification de personnes hypertendues dans une population consiste à employer un seuil de pression artérielle. Lors des répétitions des mesures, la moyenne des valeurs de pression artérielle des sujets ainsi sélectionnés lors d'une première mesure va se rapprocher de la moyenne des valeurs de la PA de l'ensemble de la population. La régression à la moyenne est d'autant plus forte que le seuil choisi est plus extrême. Ceci explique des erreurs d'analyse, telles que par exemple : l'effet du traitement est d'autant plus marqué que les taux de départ sont plus élevés.

Représentatif

Qualifie un échantillon par rapport à un caractère, si ce caractère se distribue identiquement dans l'échantillon et dans la population dont il est issu. Un échantillon tiré au sort est représentatif.

Rétrospectif

Une étude est dite rétrospective lorsque l'on y recueille des informations et des mesures concernant des événements qui se sont déroulés avant le début en pratique de l'étude.

Risque alpha ou risque de 1^{re} espèce

C'est le risque, accepté dans une étude, de conclure à une différence significative alors qu'en réalité elle n'existe pas. Ce risque est fixé arbitrairement fixé à 5% en médecine et en biologie mais d'autres valeurs peuvent se justifier.

Risque beta ou risque de 2^e espèce

C'est le risque, accepté dans une étude, de ne pas conclure à une différence significative alors qu'elle existe. Ce risque peut être diminué en augmentant les effectifs de l'étude.

Risque absolu : voir Mesures de risque

Risque attribuable : voir Mesures de risque

Risque relatif : voir Mesures de risque

Risque absolu

Risque de survenue d'un événement dans un groupe donné. Ce risque n'établit donc pas de comparaison avec un autre groupe, d'où le terme d'absolu.

Sensibilité

La sensibilité est une notion qui caractérise les qualités diagnostiques d'un signe clinique (S), d'un test clinique ou paraclinique (T). La sensibilité est la probabilité conditionnelle de la présence du signe S⁺ ou du test positif T⁺ lorsque la maladie est présente (M⁺). C'est la probabilité d'avoir un test positif lorsque l'on présente la maladie.

	M ⁺	M ⁻
S ⁺ ou T ⁺	a	B
S ⁻ ou T ⁻	c	D

La sensibilité = $[a / (a + c)]$ = Probabilité (S⁺ | M⁺) ou encore Probabilité (T⁺ | M⁺)

L'effectif a correspond aux vrais positifs, l'effectif c correspond aux faux négatifs

Signification/significativité statistique : voir principe du test statistique.

Simple – insu

Etude contrôlée dans laquelle le patient ne connaît pas dans quel groupe de traitement il est affecté (il ignore quel traitement lui est attribué). Cette méthode a pour objectif d'éviter les biais liés aux réponses du patient.

Spécificité

La spécificité est une notion qui caractérise les qualités diagnostiques d'un signe clinique (S), d'un test clinique ou paraclinique (T). La spécificité est la probabilité conditionnelle de l'absence du signe S ou de la négation du test T lorsque la maladie est absente (M⁻). C'est la probabilité d'avoir un test négatif lorsque l'on ne présente pas la maladie.

	M ⁺	M ⁻
S ⁺ ou T ⁺	a	b
S ⁻ ou T ⁻	c	d

La spécificité = $[d / (b + d)]$ = Probabilité (S⁻ | M⁻) ou encore Probabilité (T⁻ | M⁻)

L'effectif d correspond aux vrais négatifs, l'effectif b correspond aux faux positifs.

Statistiques

Méthodes essentiellement descriptives centrées sur le dénombrement.

Statistique

Ensemble de méthodes construites autour de l'organisation d'une étude, des lois de probabilités et des tests pour apporter une validation scientifique à une étude utilisant des échantillons aléatoires.

Survie

Une étude de survie est une étude dans laquelle on quantifie la distribution de la survie d'un groupe de sujet. La survie s'entend au sens large comme le temps s'écoulant entre l'entrée dans l'étude et la survenue d'un événement d'intérêt (rechute, extension, décès). Au sens strict, elle ne concerne que la survie jusqu'au décès. La méthode de Kaplan-Meier et le modèle de Cox sont les deux méthodes statistiques les plus couramment utilisées pour décrire et analyser la survie d'un groupe de patients. Le modèle de Cox permet en outre l'étude des facteurs pronostiques indépendants. Ces méthodes tiennent compte de la durée de suivi du sujet jusqu'à l'apparition de l'événement, sachant que cet événement ne sera pas forcément observé chez tous les sujets. En effet, l'étude devant pouvoir être menée en un temps raisonnable, certains patients ne présenteront pas l'événement avant la fin de l'étude.

Test diagnostique

Un test diagnostique est un examen clinique (signe, symptôme) ou paraclinique (radiographie, dosage biologique, questionnaire etc) dont le résultat permet de classer un sujet soit comme sujet sain soit comme malade. Pour évaluer l'intérêt de ce test diagnostique, on fait appel aux notions de Sensibilité, Spécificité, valeurs prédictives positives et négatives.

Tests non paramétriques

Méthodes statistiques permettant la comparaison de deux ou plusieurs groupes sans faire d'hypothèses sur la distribution des grandeurs étudiées. Ces méthodes sont basées sur des calculs de probabilités. Les tests non paramétriques doivent être utilisés lorsque les conditions nécessaires à l'utilisation des tests paramétriques ne sont pas satisfaites. Ces tests sont en général moins puissants que les tests paramétriques.

Tests paramétriques

Méthodes statistiques utilisées lorsque les lois de distribution sont connues et que celles-ci sont caractérisées par leurs paramètres principaux (moyenne, écart-type).

Valeurs manquantes

Les valeurs manquantes sont un problème très fréquent en recherche médicale. Lors de la lecture d'un article médical, il faut essayer de détecter d'éventuelles valeurs manquantes (par des effectifs variables d'une analyse à l'autre) et essayer d'évaluer l'existence de biais potentiels liés à ces données manquantes. Cela est néanmoins difficile à faire.

Valeurs prédictives positives et négatives

Ces grandeurs mesurent la performance d'une stratégie diagnostique (d'un test).

La Valeur Prédictive Positive (VPP) est la proportion de tests positifs qui correspond à de vrais malades. C'est le rapport du nombre de vrais positifs sur le nombre total de test positifs.

La VPP correspond à la probabilité pour le patient d'être malade lorsque le test est positif.

	M ⁺	M ⁻
S ⁺ ou T ⁺	a	b
S ⁻ ou T ⁻	c	d

$$VPP = a / (a+b) = \text{Probabilité} (M^+ | S^+)$$

La Valeur Prédictive Négative (VPN) est la proportion des tests négatifs qui correspond à des non-malades. C'est le rapport du nombre de vrais négatifs sur le nombre total de tests négatifs.

La VPN correspond à la probabilité pour le patient de ne pas être malade lorsque le test est négatif.

$$VPN = d / (c+d) = \text{Probabilité} (M^- | S^-)$$

Validité interne / externe

Une étude présente une validité interne si son organisation permet de répondre à ses objectifs. Une étude présente une validité externe si elle est cohérente avec les connaissances actuelles sur le sujet.

CHERCHER L'ERREUR...¹

(Lecture critique de résumé d'essai contrôlé)

Questions

1 **Diabète :**

Une étude a été menée sur 72 sujets tirés au sort dans une population de diabétiques de la maturité et ne recevant auparavant aucun traitement médicamenteux, pour savoir si un mois de traitement par un nouveau biguanide permettrait de faire diminuer la glycémie. A l'entrée dans l'essai, les moyennes de glycémie pour les sujets étaient de $8,3 \pm 2,1$ mmol. Après un mois de traitement, la glycémie moyenne était de $6,2 \pm 1,3$ mmol. La différence entre ces deux chiffres est statistiquement significative, $p < 0,01$. Cette étude montre donc l'efficacité de ce biguanide pour faire baisser la glycémie.

2 **Antihypertenseur :**

Afin de tester l'efficacité d'un nouvel antihypertenseur inhibiteur de l'enzyme de conversion, un essai croisé a été effectué en médecine de ville chez 124 patients ayant tous une hypertension artérielle essentielle d'intensité modérée. Après deux semaines de wash out, tous les patients ont eu le nouveau traitement antihypertenseur pendant six semaines. Après le cross over, tous les patients ont eu six semaines de placebo. Une analyse de variance a montré que les chiffres tensionnels diminuaient significativement plus après les six semaines de traitement par IEC qu'après les six semaines de traitement par placebo. Ces différences significatives se retrouvent aussi si on fait une comparaison de moyennes et permettent de conclure qu'il existe bien un effet antihypertenseur supérieur à celui du placebo pour ce nouvel IEC

3 **Chirurgie oesophagienne :**

Depuis Janvier 1995, le centre de chirurgie oesophagienne de l'hôpital X possède un système d'agrafage intra-oesophagien permettant de réaliser des sutures d'excellente qualité. La qualité de ce système d'agrafage a été prouvée, car durant l'année 1995, 73 patients ont pu bénéficier de ce nouvel appareillage. Il n'a été noté des complications locales post opératoires que chez 4 de ces patients. En revanche, dans le même centre chirurgical, avec les mêmes chirurgiens, durant l'année 1994 et en l'absence d'appareil d'agrafage, 83 sutures manuelles de l'oesophage avaient été réalisées et entraîné 13 complications locales post-opératoires. La différence dans la fréquence des complications est statistiquement significative en faveur de l'appareil d'agrafage (χ^2 corrigé $p < 0,02$).

¹ * Publié avec l'aimable autorisation du Professeur J.F. BERGMANN (Paris-Lariboisière)

4 **Dyspepsie :**

Afin d'étudier l'efficacité d'un nouveau médicament antidyspeptique, une étude en parallèle double aveugle randomisée contrôlée versus placebo a été menée. Tous les patients inclus avaient une dyspepsie cotée par une échelle validée, avec un score moyen de $3,2 \pm 1$ dans le groupe traité par placebo et de $3,4 \pm 0,8$ dans le groupe traité par le nouvel antispasmodique (différence non significative à l'entrée dans l'étude). Après six semaines de traitement, le score moyen de dyspepsie est passé à $1,4 \pm 0,3$ dans le groupe traité par l'antispasmodique et à $2,8 \pm 0,5$ dans le groupe traité par placebo. La différence entre ces deux scores est statistiquement significative ($p < 0,05$). Cette étude prouve l'efficacité de ce nouvel antidyspeptique comparé au placebo en terme d'amélioration d'un score de dyspepsie.

5 **Laxatif :**

Afin d'étudier l'efficacité d'un nouveau laxatif, il a été mené un essai randomisé contrôlé versus placebo. Deux groupes de 30 patients ont reçu soit le placebo, soit le laxatif. Les deux groupes étaient comparables à l'entrée dans l'essai, avec une moyenne de 2,2 selles par semaine. Après 15 jours de traitement, le nombre de selles dans le groupe placebo est passé de $2,2 \pm 0,8$ à $2,8 \pm 0,9$ (différence non significative). En revanche, dans le groupe traité par le nouveau médicament, le nombre moyen de selles est passé de $2,2 \pm 0,6$ à $3,8 \pm 0,3$ (différence significative $p < 0,01$). Ces résultats prouvent l'efficacité supérieure du nouveau laxatif par rapport au placebo en terme de nombre de selles.

6 **Radiothérapie dans le cancer du côlon :**

Afin d'évaluer l'efficacité de la radiothérapie adjuvante en association avec la chirurgie du cancer du côlon, une étude a été menée de façon randomisée et contrôlée. Tous les patients chez qui l'indication d'une colectomie pour cancer avait été posée ont été tirés au sort. La moitié des patients a eu une colectomie dans la clinique chirurgicale A; l'autre dans le centre radiochirurgical B. La moyenne de survie à 5 ans a été de 38 % chez les malades avant eu une colectomie seule et de 58 % chez les malades ayant eu colectomie plus radiothérapie. La différence entre les deux groupes est statistiquement significative. Cette étude randomisée, contrôlée et menée en parallèle montre la supériorité de l'association radiothérapie plus chirurgie par rapport à la chirurgie seule.

7 **Surfactant artificiel :**

Afin de tester l'efficacité d'un nouveau surfactant artificiel à utiliser chez les enfants prématurés en prévention de la maladie des membranes hyalines, une étude contrôlée a été menée dans une maternité parisienne. Le surfactant artificiel était instillé par un appareil spécial dont le Docteur X avait la responsabilité. Tous les enfants prématurés nés dans cette maternité lorsque le Docteur X était présent ont eu une instillation de surfactant artificiel. Les enfants nés dans cette maternité en l'absence du

Docteur X n'ont pas eu de surfactant et ont constitué le groupe contrôle. Il y a eu finalement le même nombre de patients dans les deux groupes. La survie à deux mois a été finalement de 82 % dans le groupe traité par surfactant artificiel et de 64 % dans le groupe sans surfactant. Cette différence est statistiquement significative et confirme l'efficacité de ce produit pour prévenir la mortalité des enfants prématurés.

8 **Yohimbine :**

Afin d'évaluer l'efficacité de la Yohimbine dans l'impuissance, une étude en demi cross over a été menée. Tous les patients incluables étaient dans un premier temps randomisés pour recevoir soit le placebo, soit la Yohimbine. Après trois semaines de traitement, les patients ayant reçu du placebo recevaient la Yohimbine, mais pour des raisons éthiques, les patients sous Yohimbine dans la première phase de l'étude ne recevaient pas le placebo. 22 patients ont été d'emblée traités par Yohimbine et 13 par placebo. 89 % des patients sous Yohimbine ont noté une amélioration de leur potentiel sexuel contre 49 % des patients sous placebo. Lorsque ces derniers ont été traités par la Yohimbine, 60 % d'entre eux ont noté une amélioration de leur potentiel sexuel. Dans tous les cas de figure, les résultats obtenus par la Yohimbine sont supérieurs à ceux obtenus par le placebo. Ceci confirme la supériorité de ce traitement par rapport au placebo pour le traitement des impuissances.

9 **Oxygénateur cérébral :**

On a comparé sur 2 groupes de 120 malades un oxygénateur cérébral à un placebo. Le critère d'évaluation principal était évalué par l'échelle d'appréciation clinique en gériatrie (EACG). Globalement les résultats ne montraient pas de différence significative. Cependant, si l'on éliminait les gens trop anxieux (score d'Hamilton > 18), en ne prenant en compte que les femmes (les hommes semblent moins répondeurs), la différence devient hautement significative sur la tranche d'âge 70 à 80 ans, ce qui montre bien l'activité du produit chez les gens âgés.

10 **Troubles liés au vieillissement :**

On a évalué l'activité du V. par une étude randomisée en double aveugle contre placebo, chez des patients atteints de troubles liés au vieillissement. Les critères d'évaluation déterminés à J0 J30 J90 comportaient 7 tests psychocliniques, 14 tests psychologiques, les résultats montrent la supériorité significative du V. en ce qui concerne les performances au test de Pacaud dès J 30, au trait marking test et au temps de réaction complexe à J90. Cette étude montre un effet significatif du V. en faveur d'une amélioration chiffrée des fonctions cognitives chez le sujet âgé.

11 **Extraits ribosomiaux :**

On a étudié l'effet d'injections d'extraits ribosomiaux dans la protection des poussées infectieuses chez des malades prédisposés. L'étude a consisté en l'observation comparative du "confort infectieux" de 26 patients traités depuis 2 ans par le produit, en regard des 2 années précédentes. Le nombre d'épisodes infectieux a diminué de 72 %. Le nombre de jours d'antibiotiques de 69 %, le nombre de jours d'arrêt de travail de 64 %.. L'efficacité d'un tel traitement apparaît évident.

12 **Artérite :**

Afin d'évaluer un nouveau traitement de l'artérite stade 2, on compare un extrait d'une plante africaine au tartrate d'I. Le nombre de sujets nécessaire pour mettre en évidence une amélioration de 40 % du périmètre de marche, au risque alpha de 5 %, avec une puissance de 90 %, est de 80 sujets par groupe. 236 patients répondants aux critères d'inclusion sont donc tirés au sort, pour recevoir l'un ou l'autre des traitements. Après un mois de run-in placebo (période de latence prévue au protocole avant l'introduction des principes actifs), 51 patients ayant eu une amélioration de 30% de leur périmètre de marche sont exclus de l'essai (placebo répondeurs). 101 patients ont alors été traités par la plante, et 84 par le tartrate d'I. Au terme de l'essai, 63 % des patients du groupe "plante" ont vu leur périmètre de marche amélioré de 40 %, alors qu'ils n'étaient que 47 % dans le groupe "tartrate". Cette différence est hautement significative ($p < 0,01$) en faveur de la plante.

13 **Cancer du poumon :**

Le VP 16 a été comparé au VP 17 dans le traitement du cancer épidermoïde du poumon non opérable. Le critère principal était la survie à 6 mois. Au terme de l'essai, 37 % des patients traités par VP 16 et 43 % des patients traités par VP 17 étaient encore en vie (différence NS). Mais si l'on étudie le groupe des patients qui avait une réduction de la masse tumorale de plus de 50 %, après la troisième cure, on constate que la survie est de 46 % dans le groupe VP 16, et de 59 % dans le groupe VP 17 (différence significative $p < 0,05$). On conclut donc que le VP 17 est plus efficace que le VP 16, en terme de survie à 6 mois, chez les patients répondeurs

<p>Comment seront comptés les mots lors de l'épreuve de lecture critique d'article (LCA) (CNCI-2009)</p>
--

1. Comptent comme un mot (une case) :

- un mot :
 - simple ou composé avec ou sans tiret (exemple : globulines, gamma globulines, α -trypsine...); l'article (le, la, un, l'...) associé au mot doit être dans la même case
 - une conjonction (et...);
- un nombre ou une expression chiffrée ($m \pm SD$, $p < 0,05$, $IC_{95}[a-b]$);
- un sigle (sauf s'il est attaché à un mot : Médicament[®] compte une seule case), (exemple : OBNI);
- un acronyme accepté par le CNCI (quel que soit le nombre de lettres) (exemple : Sida);
- les abréviations acceptées par le CNCI (une case par abréviation exemple Se = sensibilité = une case);
- les lettres utilisées isolément (α , β).

2. Ne comptent pas séparément (doivent donc être associés dans une case) :

- la ponctuation (. , ; ? !);
- les signes conventionnels ($>$, $<$, \geq ...);
- les guillemets;
- les parenthèses ou crochets;
- l'article (le, la, un, l'...) associé au mot;
- les numéros ou lettres d'une énumération (accompagnés ou non d'une ponctuation ou d'un tiret (ex : a, a), 1-, 1)...);
- les unités associées à un nombre (ex : 18 mg, 172 ml/min.m², 26 m/s).

3. Comptent séparément (doivent être inscrits dans des cases séparées) tous les autres cas

Résumé

De 1 à 7							
De 8 à 14							
De 15 à 21							
De 22 à 28							
De 29 à 35							
De 36 à 42							
De 43 à 49							
De 50 à 56							
De 57 à 63							
De 64 à 70							
De 71 à 77							
De 78 à 84							
De 85 à 91							
De 92 à 98							
De 99 à 105							
De 106 à 112							
De 113 à 119							
De 120 à 126							
De 127 à 133							

De 134 à 140							
De 141 à 147							
De 148 à 154							
De 155 à 161							
De 162 à 168							
De 169 à 175							
De 176 à 182							
De 183 à 189							
De 190 à 196							
De 197 à 203							
De 204 à 210							
De 211 à 217							
De 218 à 224							
De 225 à 231							
De 232 à 238							
De 239 à 245							
De 246 à 252							
De 253 à 254							

Stratégies de dépistage de l'hépatite virale C en Médecine générale

ARTICLE
ORIGINAL

Résultats d'un essai randomisé bicentrique

Françoise ROUDOT-THORAVAL (1), Elisabeth MONNET (2), Philippe MERCET (3), Anne BASTIE (4), Daniel DHUMEAUX (4),
Jean-Philippe MIGUET (3)

(1) Service de Santé Publique, Hôpital Henri-Mondor, Créteil ; (2) Service de Santé Publique et Département d'Information Médicale, (3) Service d'Hépatologie et Gastroentérologie, Hôpital Jean-Minjoz, Besançon ; (4) Service d'Hépatologie et Gastroentérologie, Hôpital Henri-Mondor, Créteil.

RÉSUMÉ

Objectifs — Améliorer le dépistage de l'hépatite virale C en médecine de ville.

Méthodes — Une action a été entreprise dans le cadre de 2 réseaux ville-hôpital. Cette action reposait sur les principes suivants : a) participation volontaire des médecins généralistes ayant bénéficié d'une formation sur l'hépatite C ; b) affectation par allocation aléatoire à deux stratégies de dépistage différentes : groupe 1 : les médecins généralistes initiaient eux-mêmes le dépistage par la recherche de facteurs de risque chez leurs malades ; groupe 2 : les médecins généralistes étaient aidés dans ce dépistage par la mise à disposition dans leur salle d'attente d'affiches et de dépliants informant sur les facteurs de risque d'hépatite C.

Résultats — Dans les 2 réseaux, 184 médecins généralistes ont été inclus dans l'essai (groupe 1, n = 90, groupe 2, n = 94). Sur une période de 15 mois, 617 sérologies ont été prescrites, 323 par les médecins généralistes du groupe 1 (âge moyen des sujets : 40 ans) et 294 par les médecins généralistes du groupe 2 (âge moyen des sujets : 44 ans) ; 489 (79,3 %) sérologies ont été effectuées (261 par le groupe 1 et 228 par le groupe 2) et 25 (5,1 %) ont été positives (15 dans le groupe 1 et 10 dans le groupe 2). Le nombre de sérologies prescrites, effectuées ou positives n'était pas statistiquement différent dans les 2 groupes. Les motifs de demande, similaires dans les 2 groupes, étaient notamment un antécédent de transfusion (27 %), une toxicomanie (6 %), une activité sérique de l'ALAT élevée ou des symptômes d'hépatite (13 %), une exposition nosocomiale (22 %). Les facteurs de risque des 25 malades dépistés étaient une toxicomanie (44 %), une transfusion avant 1991 (16 %), une activité sérique de l'ALAT élevée ou des symptômes (12 %), autres (28 %).

Conclusion — La mise en place d'une action de dépistage en médecine de ville a conduit à découvrir une sérologie positive chez 5 % des sujets testés et ce, indépendamment de la stratégie adoptée. Cependant, le faible nombre de sérologies demandées par médecin généraliste (3 sérologies en moyenne sur une période de 15 mois) suggère une absence d'exhaustivité dans l'identification des facteurs de risque et révèle la nécessité de mettre en place et d'évaluer d'autres types d'action.

Mots-clés : Hépatite C. Essai contrôlé. Dépistage. Facteurs de risque. Médecine générale.

SUMMARY

Strategies of hepatitis C virus screening in general practice. A two-centre randomized trial

Françoise ROUDOT-THORAVAL, Elisabeth MONNET, Philippe MERCET,
Anne BASTIE, Daniel DHUMEAUX, Jean-Philippe MIGUET

(Gastroenterol Clin Biol 2000;24:1037-1041)

Objectives — To improve the detection of patients infected with hepatitis C virus.

Methods — A study was undertaken in the general medicine setting in two hepatitis C networks. General practitioners volunteered and received training on hepatitis C, then were randomly assigned to one of two screening strategies : group 1 : general practitioners prescribed hepatitis C virus testing if the risk factors for HCV hepatitis C virus infection were identified during questioning of patients, group 2 : general practitioners were helped in their screening approach by posters and leaflets on the risk factors of hepatitis C virus, available in the waiting room.

Results — A total of 184 general practitioners enrolled 90 from group 1 and 94 from group 2. During a 15-month-period, 617 serologies were prescribed, 323 by general practitioners in group 1 (in patients who were an average of 40 year-old) and 294 in group 2 (in patients who were an average of 44 year-old); 489 serologies (79.3%) were actually performed (261 and 228 respectively) and 25 (5.1%) tested positive (15 and 10 respectively). The number of prescribed, performed, and positive serologies did not differ from one group to the other. The motive for hepatitis C virus screening was similar in both groups and included a history of transfusion in 27% of cases, intravenous drug use in 6%, increased ALT or symptoms compatible with hepatitis in 13%, nosocomial exposure in 22%. Risk factors in the 25 patients who were hepatitis C virus positive were drug use (44%), history of transfusion before 1991 (16%), elevated ALT or symptoms (12%), others (28%).

Conclusion — This study comparing screening strategies in general medicine, resulted in the diagnosis of hepatitis C virus infection in 5% of tested patients, regardless of the strategy. However, the fewer serologies prescribed by general practitioners (an average of 3 tests in a 15-month-period) suggests a low rate of identified risk factors in general practice, and emphasizes that other types of screening procedures should be implemented and evaluated.

Key words : Hepatitis C. Randomized trial. Screening. Risk factors. General practitioners.

Au cours des premières années qui ont suivi la mise à disposition des tests de mise en évidence des anticorps dirigés contre le virus de l'hépatite C (VHC), le diagnostic d'hépatite C a été porté prioritairement par les

médecins spécialistes chez des malades ayant une hépatopathie chronique connue, ou des sujets ayant une activité sérique élevée des aminotransférases. Ces conditions de diagnostic rendent compte des caractéristiques des malades pris en charge à cette période et notamment de la fréquence d'une cirrhose au moment du diagnostic (21 %) [1]. En 1994, les différentes études réalisées sous l'égide du Réseau National de Santé Publique permettaient d'estimer entre 500 000 et 650 000 le nombre de sujets ayant des anticorps anti-VHC, dont 80 % étaient virémiques [2] ; à cette époque, seulement 10 % de l'ensemble de la population infectée

Cette étude a été financée par le Fond de Prévoyance et d'Éducation de la Caisse Nationale d'Assurance-Maladie des Travailleurs Salariés.

Tirés à part : F. ROUDOT-THORAVAL, Service de Santé Publique, Hôpital Henri-Mondor, 94010 Créteil.

était suivi. A partir de 1994 et surtout 1996, les incitations au dépistage émanant de la Direction Générale de la Santé ont conduit à une augmentation progressive du dépistage en médecine de ville. La recommandation d'un dépistage actif, orienté sur les facteurs de risque principaux (antécédent de transfusion avant 1991 ou utilisation intra-veineuse de drogue) devait permettre le diagnostic d'un plus grand nombre de malades à un stade plus précoce de l'infection. Or, un stade histologique peu évolué au moment du diagnostic et de la mise en route du traitement est un élément de pronostic favorable, puisqu'il est associé à une meilleure réponse au traitement [3]. Pourtant, à ce jour, 30 à 40 % seulement des sujets infectés par le VHC connaîtraient leur statut sérologique, bien qu'aucune étude ne permette d'estimer correctement cette proportion. A partir d'études pilotes limitées sur une ville [4], il apparaissait que la recherche systématique de facteurs de risque d'infection par le VHC en pratique médicale courante s'avérait difficile et était de ce fait rarement faite. L'hypothèse était alors émise qu'une information donnée en salle d'attente pouvait aider les médecins généralistes en permettant d'initier le dialogue lors de la consultation et améliorer ainsi le dépistage. L'objectif de notre étude a été de comparer, dans deux sites pilotes comportant un réseau hépatite C ville-hôpital, l'efficacité de 2 stratégies de dépistage de l'hépatite C en médecine de ville : a) initiation du dépistage par le médecin, ou b) mise à disposition d'une information en salle d'attente, afin de proposer la stratégie la plus efficace, c'est-à-dire la stratégie permettant d'augmenter le nombre de sérologies prescrites et réalisées.

Méthodes

L'étude a été conduite dans 2 sites, la région sanitaire n° 6 d'Ile-de-France et le département du Doubs, selon la même méthodologie, sur une période de 15 mois (à partir d'avril 1997 dans le Doubs et de septembre 1997 dans le secteur sanitaire n° 6 d'Ile-de-France). Tous les médecins généralistes exerçant dans le secteur sanitaire couvert par le réseau étaient sollicités par courrier pour participer à une action de dépistage ; le principe de l'action était résumé dans le courrier et un coupon réponse était joint. Les médecins généralistes étaient conviés à une soirée de formation sur l'hépatite C et acceptaient, sur la base du volontariat, de participer à l'étude. Leur affectation à l'une des deux stratégies était faite par allocation aléatoire, stratifiée sur leur mode d'exercice (seul ou en groupe) et, dans le Doubs, sur le lieu d'exercice- (rural ou urbain) : groupe 1 : les médecins généralistes initiaient eux-mêmes le dépistage par la recherche de facteurs de risque chez leurs malades ; groupe 2 : les médecins généralistes étaient aidés dans ce dépistage par la mise à disposition dans leur salle d'attente d'une affiche créée pour l'étude (40 × 60 cm) et de dépliants informant les malades sur les facteurs de risque d'hépatite C. L'affiche, représentant une autruche la tête enfouie dans la terre, proposait de lire le dépliant mis à disposition et de parler de l'hépatite C avec le médecin généraliste. Le dépliant expliquait en quelques phrases simples ce qu'était l'hépatite C et ses principaux modes de contamination : transfusion et utilisation partagée de drogue en intra-veineux, mais également tatouage et hépatite C dans l'entourage.

Pour chaque prescription de sérologie, effectuée sur une feuille spéciale assurant la gratuité de l'examen au malade, le médecin remplissait un questionnaire. Celui-ci était constitué de trois feuillets autocopiants et comportait : âge et sexe du malade, personne à l'origine de la demande (malade ou médecin), motif du dépistage (transfusion avant 1991, toxicomanie intraveineuse ou autre motif individuel de dépistage), date de la prescription, résultat de la sérologie, nom du médecin spécialiste ou service de spécialité auquel était adressé le malade. Le premier feuillet était envoyé au responsable de l'étude de chaque réseau dès prescription de la sérologie avec les renseignements concernant le malade. Le médecin généraliste expédiait secondairement, dès réception du résultat de la sérologie, le deuxième feuillet comportant le résultat de la sérologie, ou la mention « sérologie non reçue à un mois » et, éventuellement le nom du médecin spécialiste, ou le service de spécialité auquel était adressé le sujet ayant des anticorps anti-VHC positifs.

Une lettre informant sur les objectifs et le déroulement de l'étude était remise aux malades susceptibles d'être dépistés, conformément aux recommandations de la Commission Nationale Informatique et Liberté.

Analyse statistique — Les données ont été centralisées dans un premier temps au niveau de chaque site, puis regroupées pour une analyse globale. Deux unités d'analyse ont été prises en compte : le médecin d'une part, la prescription d'une sérologie d'autre part. Une analyse univariée a permis de comparer : a) les deux groupes de médecins en ce qui concerne leur mode d'exercice et leur prescription ; b) les prescriptions de sérologie selon la stratégie allouée au médecin prescripteur. Le test du χ^2 a été utilisé pour la comparaison de variables qualitatives et le test non paramétrique de Mann-Whitney pour les variables quantitatives. Une analyse multivariée par régression logistique pas à pas a été utilisée pour étudier : a) les facteurs indépendamment liés à la réalisation ou non de la sérologie prescrite ; b) les facteurs liés à un résultat positif de la sérologie virale C.

Résultats

Cent quatre-vingt quatre médecins généralistes ont été volontaires pour participer à l'étude, 119 dans le Doubs (parmi 512) et 65 dans le secteur sanitaire n° 6 d'Ile-de-France (parmi 936). La stratégie classique a été allouée à 90 médecins et la stratégie de mise à disposition d'une information à 94 médecins. La comparaison de ces deux groupes est résumée dans le tableau I. Parmi les médecins généralistes ayant participé à l'étude, seuls 56 % ont effectivement prescrit au moins une sérologie, pourcentage non statistiquement différent dans les 2 groupes. Le nombre moyen de sérologies demandées sur la période d'étude était de 3,4 par médecin généraliste inclus et de 6,0 par médecin généraliste ayant demandé au moins une sérologie, soit environ 1 demande de sérologie par trimestre et par médecin généraliste participant. Ces résultats n'étaient pas statistiquement différents dans les 2 groupes de médecins généralistes. Ils n'étaient pas différents quand on prenait en compte le site (Ile-de-France ou Doubs) ou le mode d'exercice des médecins généralistes (données non présentées). Cette similitude de résultats entre les 2 sites nous a autorisé à analyser par la suite l'ensemble des données.

Sur la période d'étude de 15 mois, 617 sérologies ont été prescrites chez 617 sujets âgés de 6 mois à 85 ans (607 adultes de plus de 15 ans), 328 hommes et 289 femmes ; 323 sérologies l'ont été par les médecins généralistes du groupe 1 (sérologie initiée par les médecins généralistes) et 294 par les médecins généralistes du groupe 2 (information mise à disposition). Les principaux résultats sont résumés dans le tableau II. La sérologie était prescrite à la demande du malade dans 19,5 % des cas dans le groupe 1 et dans 35,7 % des cas dans le groupe 2 ($P < 0,001$). Les malades étaient un peu moins âgés dans le groupe 1 (40 ± 15 ans) que dans le groupe 2 (44 ± 15 ans, $P = 0,002$), mais le sexe-ratio était similaire. Les motifs du dépistage n'étaient globalement pas différents dans les 2 groupes et étaient notamment un antécédent de transfusion avant 1991 dans 27 % des cas, une toxicomanie intra-veineuse dans 6 % des cas, une activité sérique de l'ALAT élevée ou des symptômes d'hépatite dans 13 % des cas, une exposition nosocomiale dans 22 % des cas. En revanche, les motifs de dépistage étaient significativement plus souvent multiples dans le groupe 2 (22,7 %) que dans le groupe 1 (12,0 %, $P = 0,001$). Parmi les sérologies effectuées (environ 80 % des sérologies prescrites), 5,1 % se sont révélées positives, avec une fréquence similaire dans les 2 groupes. Ainsi, des anticorps anti-VHC ont été mis en évidence chez 25 malades pendant la durée de l'étude. Les facteurs de risque ayant conduit au dépistage étaient dans 11 cas une toxicomanie intra-veineuse, dans 4 cas un antécédent de transfusion avant 1991, dans 3 cas une élévation de l'activité sérique de l'ALAT ou des symptômes compatibles avec une hépatite, dans 1 cas un tatouage isolé, dans 6 cas d'autres

Tableau I. – Comparaison des deux groupes de médecins généralistes selon le mode d'exercice et la prescription de sérologies virales C.
Comparison of two groups of General Practitioners according to their type of practice and prescription of hepatitis C virus testing.

	Groupe 1 Sérologies initiées par le MG (n = 90)	Groupe 2 Information mise à disposition (n = 94)	
Mode d'exercice			
Rural	32 (36 %)	28 (30 %)	NS
En groupe	43 (48 %)	45 (48 %)	NS
Prescripteurs de sérologie	55 (61 %)	48 (51 %)	NS
Nombre de prescriptions/MG (min-max)	3,6 (0-27)	3,1 (0-33)	NS
Nombre de prescriptions/MG participant ^a (min-max)	5,9 (1-27)	6,1 (1-33)	NS
Nombre de sérologies positives/MG (min-max)	0,27 (0-2)	0,21 (0-2)	NS

MG : médecin généraliste ; n : nombre ; NS : différence non significative. ^a Ayant demandé au moins une sérologie.

causes : hépatite C dans l'entourage (2 cas), séjours en pays à forte prévalence du VHC (2 cas), autres pratiques à risque (2 cas) (tableau III). La répartition des facteurs de risque n'était pas différente dans les 2 groupes.

En analyse multivariée, l'âge supérieur à 60 ans et le sexe masculin étaient associés positivement à la réalisation de la sérologie prescrite ; en revanche, l'existence d'une toxicomanie ou d'un tatouage était associée à une chance 10 fois moindre de réalisation de la sérologie que la constatation d'une élévation de l'activité sérique de l'ALAT ou de symptômes (tableau IV). Le seul facteur fortement associé à la positivité de la sérologie virale C était la présence d'une toxicomanie intra-veineuse comme facteur de risque (tableau V).

Discussion

L'action de dépistage en médecine de ville, orientée sur la recherche de facteurs de risque et comparant deux stratégies de

dépistage, initiation du dépistage par le médecin généraliste ou information mise à disposition, n'a pas permis de mettre en évidence de supériorité d'une stratégie de dépistage sur une autre. En effet, aussi bien le nombre de sérologies prescrites (3,6 versus 3,1 par médecin généraliste) que le pourcentage de sérologies positives parmi les tests effectués (5,7 versus 4,4 %) étaient comparables dans les 2 groupes. L'existence d'une information affichée chez le médecin généraliste a été associée à une demande plus fréquente de sérologie par le malade lui-même (35 % versus 19,5 %), montrant bien le rôle positif de l'information proposée aux malades. Toutefois, cet accroissement de demande faite par le malade lui-même n'a pas conduit à un dépistage plus fréquent du VHC dans ce groupe ; cela pose la question d'une information insuffisante, ne permettant pas au malade de bien se situer par rapport au risque d'infection par le VHC. Il est intéressant de noter à cet égard que dans ce groupe, les motifs de demande étaient plus souvent multiples et autres que les deux grands facteurs de risque justifiant le dépistage, un antécédent de transfusion avant 1991 et une toxicomanie, mais sans relation avec les autres facteurs de risque évoqués sur le dépliant (tatouage, hépatite C dans l'entourage). Ces résultats

Tableau II. – Comparaisons des caractéristiques liées aux prescriptions en fonction de la stratégie allouée aux médecins généralistes.
Comparisons of the prescriptions according to the strategy assigned to the General Practitioners.

	Groupe 1 sérologies initiées par le MG (n = 323)	Groupe 2 information mise à disposition (n = 294)	P
Sérologie demandée par le malade	63 (19,5 %)	105 (35,7 %)	0,001
Age moyen des malades (1ds)	40,1 (15,2)	43,9 (15,1)	0,002
Sexe : % hommes	56,0	50,0	NS
Motif de la sérologie			
Antécédent de transfusion	82 (25,4 %)	82 (27,9 %)	
Toxicomanie	34 (10,5 %)	17 (5,8 %)	
Tatouage	68 (21,1 %)	65 (22,1 %)	NS
ALAT élevées ou symptômes	45 (13,9 %)	34 (11,6 %)	
Exposition nosocomiale	64 (19,8 %)	71 (24,1 %)	
Autres	75 (23,2 %)	89 (30,3 %)	
Nombre de FR identifiés			
1 FR	276 (85,4 %)	216 (73,5 %)	
2 FR	34 (10,5 %)	56 (19,0 %)	0,001
3 FR et plus	5 (1,5 %)	11 (3,7 %)	
Nombre d'examens effectués	261 (80,8 %)	228 (77,6 %)	NS
Nombre de sérologies positives	15 (5,7 %)	10 (4,4 %)	NS

FR : facteur de risque ; NS : non significatif ; MG : médecin généraliste ; n : nombre.

Tableau III. – Motifs de prescription de la sérologie virale C chez les 25 malades positifs et proportion de sérologies positives parmi les sérologies réalisées, selon le motif de prescription

Motive for prescribing hepatitis C virus testing in the 25 individuals who tested positive for hepatitis C virus antibodies and proportion of hepatitis C virus positive tests according to the motive for prescribing testing.

Motif de prescription	Fréquence du facteur	Prévalence du VHC pour le facteur
Toxicomanie	11 (44 %)	11/24 (45,8 %)
Transfusion avant 1991	4 (16 %)	4/141 (2,8 %)
ALAT élevées ou symptômes	3 (12 %)	3/74 (4,1 %)
Tatouage	1 (4 %)	1/77 (1,3 %)
Autres ^a	6 (24 %)	

VHC : virus de l'hépatite C. ^a VHC dans l'entourage : 2 cas, séjour en pays de forte prévalence : 2 cas, autres pratiques à risque : 2 cas.

montrent bien l'absence de pertinence et de rentabilité du dépistage, dès que l'on élargit ses critères [5].

La mise en place de l'action de dépistage dans les 2 sites a permis une sensibilisation d'environ 20 % des médecins généralistes du Doubs et 10 % seulement des médecins généralistes du secteur sanitaire 6 d'Ile-de-France. Les sollicitations diverses et nombreuses que subissent l'ensemble des médecins généralistes, notamment en région urbaine comme l'Ile-de-France, pour participer à des actions ou des formations sur diverses pathologies, peut expliquer la faible proportion de médecins généralistes réunis sur une pathologie donnée, l'hépatite C, même si la majorité des médecins généralistes français interrogés déclarent souhaitable le dépistage de l'hépatite C [6]. Dans cette population sélectionnée de médecins généralistes volontaires pour participer à l'action, l'absence de supériorité de la stratégie « information mise à disposition » pourrait donc être liée à une performance de dépistage particulière des médecins généralistes initiant eux-mêmes le dépistage. Cependant, les faibles résultats observés vont à l'encontre d'une telle hypothèse.

L'action de dépistage, qui a été menée, a permis la mise en évidence d'une sérologie virale C positive chez 5 % des malades

ayant effectué la sérologie prescrite. Cette prévalence, comparée au 1 % de prévalence du VHC dans la population générale, montre la rentabilité du dépistage orienté. Nos résultats sont à rapprocher des résultats de 3 à 8 % obtenus lors d'actions de dépistage en médecine de ville menées dans d'autres régions [7-9]. Dans 60 % des cas dépistés, la prescription était en rapport avec une toxicomanie ou une transfusion avant 1991, qui sont les deux seuls facteurs de risque pour lesquels un dépistage systématique a été recommandé par la conférence de consensus française [10]. On observe d'ailleurs dans ces groupes une prévalence du VHC de 45,8 % et 2,8 % respectivement pour la toxicomanie et pour les antécédents de transfusion. La toxicomanie reste à l'évidence la conduite la plus liée à l'existence d'une séropositivité VHC. La constatation d'une élévation de l'activité sérique de l'ALAT demandée pour d'autres motifs ou de symptômes compatibles avec une hépatite est associée à une prévalence du VHC de près 4,1 %, ce qui correspond aux données de la littérature [4, 11-13]. En revanche, un tatouage isolé ne constitue pas un excès de risque d'infection par le VHC, la prévalence dans ce groupe étant semblable dans notre échantillon (1,3 %) à celle de la population générale. Les autres facteurs de risque ne rendent compte que d'un quart des malades dépistés.

Tableau IV. – Facteurs associés à la réalisation de la sérologie virale C prescrite : résultats de l'analyse multivariée.

Factors associated with performing hepatitis C virus testing: results of multivariate analysis.

	Sérologies faites (%)	OR	IC 95 %	P
Age				
< 40 ans	212/305 (69,5)	1 ^a		
40-59 ans	173/212 (81,6)	1,6	0,9-3,1	0,14
60 ans et plus	90/100 (90,0)	3,8	1,1-13,3	0,03
Sexe				
féminin	224/289 (77,5)	1 ^a		
masculin	264/328 (80,5)	2,0	1,1-3,3	0,03
FR				
ALAT > N ou symptômes	74/79 (93,7)	1 ^a		
toxicomanie	24/51 (47,1)	0,04	0,005-0,3	0,005
tatouage	77/110 (70,0)	0,06	0,007-0,4	0,008
transfusion	141/164 (86,0)	0,2	0,03-1,7	0,14
exposition nosocomiale	63/76 (82,9)	0,2	0,02-1,7	0,09
autres	110/137 (80,3)	0,2	0,02-1,3	0,13

N : limite supérieure de la normale ; OR : odd ratio ; IC : intervalle de confiance. ^aCatégorie de référence.

Tableau V. – Facteurs associés à la positivité de la sérologie virale C : résultats de l'analyse multivariée

Factors associated with hepatitis C virus positive tests: results of multivariate analysis.

	Sérologies positives (%)	OR	IC 95 %	P
Age				
< 40 ans	14/212 (6,6)	1 ^a		
40-59 ans	8/173 (4,6)	1,4	0,5-4	0,51
60 ans et +	3/90 (3,3)	1,7	0,4-7,4	0,48
Sexe				
féminin	10/224 (4,5)	1 ^a		
masculin	15/264 (5,7)	1,0	0,4-2,7	0,93
FR				
ALAT > N ou symptômes	3/74 (4,1)	1 ^a		
toxicomanie	11/24 (45,8)	19,8	4,4-89,6	0,0001
tatouage	1/77 (1,3)	0,3	0,03-2,8	0,28
transfusion	4/141 (2,8)	0,6	0,1-2,7	0,44
exposition nosocomiale	0/63 (0)	0,01	0,00-4,10 ⁷	0,56
autres	6/110 (5,5)	1,3	0,3-5,4	0,74

FR : facteur de risque ; OR : odd ratio ; IC : intervalle de confiance. ^aCatégorie de référence.

Le nombre de 25 malades dépistés sur une période de 15 mois par près de 200 médecins généralistes, dans deux sites comportant au total plus d'un million d'habitants adultes peut paraître très faible. Le nombre de sérologies demandées par médecin généraliste est également très faible, en moyenne 6 par médecin généraliste ayant demandé au moins une sérologie, soit une moyenne d'une sérologie tous les deux mois. Ces résultats suggèrent fortement une absence d'exhaustivité du dépistage orienté sur facteurs de risque. Ainsi, le petit nombre de malades (164) ayant eu une demande de sérologie pour antécédent de transfusion est probablement très inférieur au nombre de malades qui, dans la clientèle des médecins généralistes de l'étude, ont été effectivement transfusés. Cependant, dans l'hypothèse où la proportion de médecins généralistes ayant demandé au moins une sérologie (environ 10 %) couvrent un dixième de la population adulte, l'incidence des cas nouvellement dépistés au cours de notre action peut être estimée à environ 25 pour 100 000, comparable avec celle relevée dans les différents registres d'hépatite [14]. Ainsi, malgré une action spécifique ayant donné lieu à un volontariat de participation, l'incidence des nouveaux cas diagnostiqués n'est pas sensiblement améliorée par cette action spécifique. Ces résultats peuvent refléter la difficulté du dépistage en médecine de ville, ou la difficulté d'une action soutenue dans le temps, comme en témoigne la nette diminution du nombre de sérologies prescrites entre le premier et le dernier trimestre de l'étude (nombre divisé par 7 dans le secteur sanitaire 6 d'Ile-de-France, résultats non montrés). Ainsi, la recherche d'autres types d'action, peut-être plus concentrées dans le temps, et/ou élargissant les cibles de l'information, pourraient être nécessaires pour dépister les sujets dont la séropositivité VHC reste inconnue.

En conclusion, le dépistage de l'infection par le VHC en médecine de ville, orienté sur l'existence de facteurs de risque permet d'obtenir une prévalence de 5 % parmi les consultants testés mais n'est pas amélioré par des stratégies visant à aider les médecins dans leur travail de dépistage. Ainsi, le faible nombre de malades dépistés rend compte d'une imperfection du dépistage tel qu'il est actuellement effectué. Ces résultats révèlent le besoin de mettre en place et d'évaluer d'autres types d'action de dépistage si l'on veut atteindre les objectifs fixés par les responsables de la Santé Publique : 80 % des malades devront connaître leur statut viral C à la fin de l'année 2000.

REMERCIEMENTS - Cette étude a été financée par le Fonds de Prévoyance et d'Education de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés.

Nous remercions, de plus, les Laboratoires Schering-Plough pour l'aide apportée à la réalisation de notre action de dépistage, en ce qui concerne la fourniture gracieuse des dépliants d'information aux malades et le remboursement des tests de dépistage aux laboratoires de biologie médicale.

RÉFÉRENCES

1. Roudot-Thoraval F, Bastie A, Pawlotsky JM, Dhumeaux D and the Study Group for the Prevalence and the Epidemiology of Hepatitis C. Epidemiological factors affecting the severity of hepatitis C virus-related liver disease : a French survey of 6,664 patients. *Hepatology* 1997;26:485-90.
2. Desenclos JC, Dubois F, Couturier E, Pillonel J, Roudot-Thoraval F, Guignard E, et al. Estimation du nombre de sujets infectés par le VHC en France, 1994-1995. *BEH* 1996;5:22-3.
3. Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, et al. Randomised trial of interferon α 2b plus ribavirin for 48 weeks or for 24 weeks versus interferon α α 2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. *Lancet* 1998;352:1426-32.
4. Roudot-Thoraval F, Tuffery S, Deforges L, Pawlotsky JM, Saint-Marc Girardin MF, Dhumeaux D. Dépistage des hépatites C en médecine de ville par une demande orientée des transaminases (résumé). *Gastroenterol Clin Biol* 1993;17:685.
5. Capron D, Bensoussan T, Darchis JP, Barbare JC, Butel J, Bental A, et al. Hepatitis C virus infection risk factors in patients admitted in hospital emergency departments in Picardy. Value of oriented screening based on recommendations of the 'Direction Générale de la Santé'. *Eur J Gastroenterol Hepatol* 1999;11:643-8.
6. Babany G, Bourlière M, Chevalier H, Chousterman M, Couzigou P, Desmorat H, et al. Les médecins généralistes souhaitent-ils prendre en charge l'hépatite chronique virale C et participer aux réseaux ville-hôpital ? *Gastroenterol Clin Biol* 1999;23:1289-95.
7. Frère T, Verneau A, Besson I, De Baysier L, Dupuychaffray J-P, Kitmacher P, et al. Prise en charge de l'infection par le virus de l'hépatite C dans la région Poitou-Charentes. *Gastroenterol Clin Biol* 1999;23:887-91.
8. Bonnaud G, Crémieux A-C, Vuagnat A, Leroy C, Lucet JC, Bouchaud O, et al. Evaluation de différentes stratégies de dépistage de l'hépatite C en médecine générale et dans un centre de dépistage anonyme et gratuit (CDAG) (résumé). *Gastroenterol Clin Biol* 1999;23:A98.
9. Altman C, Lesieur A, Meyer L, Dunbavand A, de La Selle P, Zrihen E, et al. Dépistage de l'hépatite virale C en médecine générale (résumé). *Gastroenterol Clin Biol* 1999;23:A170.
10. Conférence de Consensus. Hépatite C : dépistage et traitement. *Gastroenterol Clin Biol* 1997;20:S202-S211.
11. Jacques D, Vassal D. Intérêt d'une stratégie de dépistage cible de l'hépatite C dans un centre d'examen de santé (résumé). *Gastroenterol Clin Biol* 1999;23:A171.
12. Dubois F, François M, Mariotte N, Caces E, Vol S, Roingard P, Barin F, et al. Serum alanine aminotransferase measurement as a guide to selective testing for hepatitis C during medical check-up. *J Hepatol* 1994;21:837-41.
13. Desenclos JC, Dubois F, Mariotte N, Goudeau A. Analyse des stratégies de dépistage orienté de l'infection par le virus de l'hépatite C. *Gastroenterol Clin Biol* 1997;20:S25-32.
14. Minello A, Boschi F, Harb M, Milan C, Faivre J, Hillon P et le Réseau Bourguignon des Hépatites C. Mise en place d'un registre des hépatites virales B et C dans le département de la Côte-d'Or. *Méthodologie, premiers résultats. Gastroenterol Clin Biol* 1998;22:766-71.

**Stratégie de dépistage de l'hépatite C en médecine générale.
Résultats d'un essai randomisé bicentrique**

Question 1 :

Pour cet essai randomisé visant à démontrer l'efficacité d'une intervention précisez :

- a) la nature précise de l'intervention qu'on cherche à évaluer
- b) le ou les critères de jugements utilisés
- c) pourquoi y a-t-il « deux niveaux d'analyse » ?

Question 2 :

Pourquoi a-t-on stratifié la randomisation sur le mode d'exercice ?

Question 3 :

Les auteurs considèrent que le résultat de l'essai est non concluant : quelles explications vous semblent possibles pour expliquer ce résultat ?

Question 4 :

Les résultats obtenus dans cet essai vous paraissent-ils applicables à l'ensemble des médecins généralistes en France ? Justifiez votre réponse.

Question 5 :

Les résultats obtenus dans cet essai vous paraissent-ils applicables à l'ensemble de la population française ? Justifiez votre réponse.

Question 6 :

Peut-on comparer la clientèle des deux groupes ? Justifiez votre réponse.

Question 7 :

Le titre de cet article vous semble-t-il adapté à son contenu ? Pouvez-vous en proposer un autre ?

ÉNOXAPARINE COMPARÉE À UN PLACEBO DANS LA PRÉVENTION DE LA MALADIE
THROMBO-EMBOLIQUE VEINEUSE CHEZ DES PATIENTS PRÉSENTANT UNE
PATHOLOGIE MÉDICALE AIGUË

LA MALADIE thrombo-embolique veineuse est une découverte fréquente à l'autopsie chez des patients décédés à l'hôpital sous traitement médical.^{1,2} Bien que la fréquence de la maladie thrombo-embolique veineuse chez ces patients n'ait pas été établie, elle est rapportée comme étant presque aussi élevée que chez les patients chirurgicaux à risque thrombo-embolique modéré.^{3,4}

Il existe de nombreuses données soutenant à la fois le bénéfice clinique et le rapport coût-efficacité de la thrombo-prophylaxie systématique chez les patients chirurgicaux,^{3,9} mais l'utilisation de cette approche chez des patients de médecine générale reste controversée. Pour certains groupes de patients — tels que ceux qui ont un accident vasculaire cérébral avec paralysie ou un infarctus du myocarde, chez lesquels l'incidence de la maladie thrombo-embolique veineuse varie de 30 à 75 % — l'utilisation prophylactique de l'héparine est recommandée.^{3,4,10} Pour les autres patients médicaux hospitalisés, la situation est moins claire en raison de l'hétérogénéité des schémas des études disponibles, des méthodes différentes utilisées pour diagnostiquer la maladie thrombo-embolique veineuse et, surtout, en raison de l'hétérogénéité des patients.³ De plus, le risque de maladie thrombo-embolique veineuse peut varier en fonction de la présence de facteurs de risque intrinsèques¹¹⁻¹³ et peut ainsi expliquer ces résultats contradictoires.¹⁴⁻²⁴

La fréquence de la maladie thrombo-embolique veineuse n'étant pas connue chez les patients hospitalisés en médecine et les preuves de l'efficacité de la prophylaxie systématique étant insuffisantes, nous avons réalisé une étude randomisée, en double-aveugle, contrôlée contre placebo chez ce type de patients avec deux objectifs : déterminer la fréquence des thromboses veineuses profondes et des embolies pulmonaires et déterminer l'efficacité et la tolérance de deux posologies de traitement d'héparine de bas poids moléculaire dans la prévention des thromboses veineuses profondes et des embolies pulmonaires.

MÉTHODES

Patients

Étaient considérés pour inclusion dans l'étude, les patients hospitalisés en médecine, âgés de plus de 40 ans, dont le séjour envisagé était de plus de six jours, non immobilisés depuis plus de trois jours. Pour être éligibles, les patients devaient avoir soit une insuffisance cardiaque congestive (classe III ou IV de la New York Heart Association), soit une insuffisance respiratoire aiguë ne nécessitant pas d'assistance ventilatoire ou une des pathologies médicales suivantes associée à au moins un facteur additionnel de risque thrombo-embolique veineux : infection aiguë sans choc septique, affections rhumatismales aiguës incluant les douleurs lombaires ou les sciatiques ou les compressions vertébrales (conséquences d'une ostéoporose ou d'une tumeur), les arthrites aiguës ou les épisodes aigus d'arthrite rhumatoïde des membres inférieurs ou les poussées de maladie inflammatoire intestinale. Les facteurs de risque additionnels étaient un âge de plus de 75 ans, un cancer, un antécédent thrombo-embolique veineux, l'obésité (index de masse corporelle [poids en kilogrammes divisé par la taille en mètres au carré], ≥ 30 pour les hommes et 28,6 pour les femmes), les varices, un traitement hormonal (par antiandrogène ou estrogène, à l'exception des traitements hormonaux substitutifs de la ménopause) et l'insuffisance cardiaque ou respiratoire chronique.

Les femmes en âge de procréer étaient exclues en cas de grossesse, d'allaitement au sein ou en l'absence de contraception ; les autres raisons d'exclusion étaient les accidents vasculaires cérébraux ou les interventions chirurgicales majeures au cours des trois mois précédents, les contre-indications à l'utilisation des produits de contraste iodés, les thrombophilies connues, une créatinémie supérieure à 1,7 mg/dl (150 $\mu\text{mol/l}$), une intubation, la présence d'une infection à VIH, une pression artérielle non contrôlée (PAS > 200 mm Hg ou PAD > 120 mm Hg ou les deux), un ulcère peptique évolutif, une endocardite bactérienne ou d'autres pathologies pouvant augmenter le risque d'hémorragie telles qu'une hypersensibilité à l'héparine ou une thrombopénie induite par l'héparine, un nombre de plaquettes < 100 000/mm³, un allongement du temps de céphaline activée, un taux de prothrombine inférieur à 50 % ou un INR > 1,2. De plus, les patients qui nécessitaient un traitement anticoagulant et ceux qui avaient reçu un anticoagulant pendant plus de 48 heures étaient exclus.

Schéma de l'étude

La randomisation était réalisée de façon centralisée. Les patients étaient randomisés en vue de recevoir 20 mg ou 40 mg d'énoxaparine (Lovenox, Clexane ou Klexane, Rhône-Poulenc Rorer, Antony, France) ou un placebo par voie sous-cutanée une fois par jour, dans les 24 heures suivant la randomisation. Le traitement était prévu pour une durée de 6 à 14 jours à l'hôpital. Les patients étaient alors suivis directement ou par téléphone entre les jours 83 et 110. Lors de ce suivi, les patients devaient rapporter tous symptômes ou signes de maladie thrombo-embolique veineuse ou tout autre événement clinique survenu après la fin du traitement.

L'étude était réalisée selon les principes éthiques de la Déclaration d'Helsinki et selon les réglementations locales. Le protocole était approuvé par un comité d'éthique indépendant ou éventuellement par le comité d'éthique des centres et un consentement éclairé écrit était obtenu de tous les patients avant randomisation.

Traitements

Tous les médicaments de l'étude étaient préparés dans des seringues à dose unique préremplies contenant soit 40 mg d'énoxaparine dans 0,2 ml d'eau pour préparations injectables (à une concentration de 200 mg/ml équivalente à 20 000 UI anti Xa/ml), soit 20 mg d'énoxaparine dans 0,2 ml/1 (100 mg/ml, équivalent à 10 000 UI anti Xa/ml) soit un placebo (0,2 ml de sérum physiologique).

Pendant toute la durée du traitement, les injections intramusculaires, les traitements avec des produits néphrotoxiques, en particulier les antibiotiques néphrotoxiques, n'étaient pas autorisés. Les autres traitements, bandages élastiques ou bas de contention et kinésithérapie étaient utilisés selon les usages habituels de chaque centre. Les centres devaient éviter si possible d'administrer des anti-inflammatoires non stéroïdiens.

Critères de jugement

Le critère de jugement principal d'efficacité était la survenue d'un événement thrombo-embolique (définie par une thrombose veineuse profonde, une embolie pulmonaire ou les deux) entre les jours 1 et 14. Le critère de jugement secondaire d'efficacité était la survenue d'un événement thrombo-embolique entre les jours 1 à 110. Une recherche systématique des TVP était effectuée par une phlébographie ascendante des membres inférieurs, réalisée entre J6 et J14 ou plus précocement si une thrombose était suspectée cliniquement. Si la phlébographie était irréalisable, un écho-Doppler veineux était fait. En cas de suspicion, le diagnostic d'embolie pulmonaire était confirmé par une scintigraphie pulmonaire de haute probabilité, par une angiographie pulmonaire, par un scanner hélicoïdal ou à l'autopsie.

L'incidence des décès, hémorragies mineures ou majeures, thrombopénies ou de tout autre événement intercurrent ainsi que les anomalies biologiques était également examinée. L'hémorragie était classée comme majeure si elle était cliniquement patente et associée à la nécessité d'une transfusion de deux unités ou plus de globules rouges ou de sang total ou à une diminution de 2,0 g/dl ou plus de la concentration d'hémoglobine par rapport aux valeurs initiales, si le saignement était rétroperitonéal, intracranien ou létal. Les hémorragies étaient définies comme mineures si elles étaient cliniquement patentes mais ne répondaient pas aux critères d'hémorragie majeure. Le site d'injection était évalué quotidiennement pour les réactions locales (hématomes supérieurs à 5 cm de diamètre). Une numération formule sanguine était pratiquée avant le traitement et tous les trois jours par la suite. Une thrombopénie était définie par une diminution du nombre de plaquettes d'au moins 30 % par rapport à la valeur initiale ou à moins de 100 000/mm³ et la thrombopénie était considérée comme sévère si le nombre de plaquettes était inférieur à 50 000/mm³.

Les événements étaient revus par deux comités indépendants dont les membres ignoraient les traitements alloués aux patients. Deux radiologues revoyaient toutes les phlébographies et les angiographies et deux spécialistes de médecine nucléaire revoyaient les scintigraphies pulmonaires. Tout désaccord était résolu par consensus. Le résultat de ces relectures était transmis au comité des événements critiques qui validait tous les événements cliniques.

Analyse statistique

Afin de maintenir un taux de significativité bilatérale globale de 0,05 dans l'analyse du critère principal, le taux de significativité nominal de l'une des analyses intermédiaires d'efficacité a été ajusté selon la méthode d'O'Brien-Fleming et dans deux analyses de comparaison avec le placebo selon la méthode de Bonferroni. Avec l'hypothèse d'une incidence d'événement veineux thrombo-embolique de 15 % dans le groupe placebo et de 6 % dans l'un des groupes énoxaparine et une puissance globale de 90 %, nous avions besoin de 284 patients dans chaque groupe (au total, 852 patients). Le nombre cible de patients inclus était de 1 020 patients, nombre qui tenait compte d'une incidence de patients non évaluables en efficacité allant jusqu'à 20 %. En avril 1998, après l'inclusion de 750 patients, le Comité Directeur augmentait

TABEAU 1. NOMBRES DE PATIENTS INCLUS DANS LES ANALYSES ET MOTIFS D'EXCLUSION.*

TYPE D'ANALYSE	TOTAL (N = 1102)	PLACEBO (N = 371)	ÉNOXAPARINE	ÉNOXAPARINE
			20 mg (N = 364)	40 mg (N = 367)
	nombre (%)			
Analyse du critère principal (jours 1-14)				
Évalue	866 (78,6)	288 (77,6)	287 (78,8)	291 (79,3)
Non évalué	236 (21,4)	83 (22,4)	77 (21,2)	76 (20,7)
Décès	28 (2,5)	11 (3,0)	10 (2,7)	7 (1,9)
Refus du patient	62 (5,6)	22 (5,9)	20 (5,5)	20 (5,4)
Décision de l'investigateur	58 (5,3)	22 (5,9)	17 (4,7)	19 (5,2)
Phlébographie non réalisable	12 (1,1)	3 (0,8)	5 (1,4)	4 (1,1)
Phlébographie non évaluable	72 (6,5)	22 (5,9)	24 (6,6)	26 (7,1)
Pour raison inconnue, phlébo- graphie non réalisée	4 (0,4)	3 (0,8)	1 (0,3)	0
Analyse du critère secondaire (jours 1-110)				
Évalue	798 (72,4)	263 (70,9)	263 (72,3)	272 (74,1)
Non évalué	71 (6,4)	26 (7,0)	25 (6,9)	20 (5,4)
Décès	61 (5,5)	23 (6,2)	20 (5,5)	18 (4,9)
Perdu de vue ou consultation non prévue avant le 90 ^e jour	10 (0,9)	3 (0,8)	5 (1,4)	2 (0,5)
Traitement avec au moins une dose du médicament à l'étude	1073 (97,4)	362 (97,6)	351 (96,4)	360 (98,1)

* Des données pour l'évaluation entre J14 et J110 sont manquantes chez 71 patients, mais 3 patients sans données d'efficacité au 14^e jour ont présenté une maladie thrombo-embolique veineuse symptomatique au cours du suivi.

le nombre cible à 1 100 pour être certain qu'un nombre de 852 patients pourrait être étudié.

L'analyse du critère principal incluait les données de tous les patients selon le principe de l'intention de traiter. L'analyse des événements indésirables incluait les données des patients ayant reçu au moins une dose du médicament étudié. Si les résultats des phlébographies ou des écho-Doppler obtenus entre les jours 6 et 14 n'étaient pas conclusifs, ils n'étaient pas analysés.

L'analyse du critère de jugement principal rendait compte de l'éventualité que les risques absolus d'événement veineux thrombo-embolique chez les patients asymptomatiques puissent différer en fonction du type d'examen des patients, phlébographie systématique ou écho-Doppler, en raison des sensibilités différentes des deux méthodes. Nous avons supposé que les deux méthodes détecteraient les événements ayant des risques relatifs similaires dans les deux groupes. Les risques relatifs d'événement veineux thrombo-embolique ont été calculés pour les patients examinés avec chaque méthode et les risques relatifs ont été ensuite combinés en utilisant le test du chi-2 de Mantel-Haenszel.²⁶ L'intervalle de confiance à 97,6 % du risque relatif commun était calculé avec une approximation normale pour une distribution binomiale. Une valeur de p bilatérale inférieure à 0,02 a été considérée comme statistiquement significative pour le critère de jugement principal. Un test du chi-2 bilatéral ou un test exact de Fisher (lorsqu'adapté) ont été utilisés pour les variables qualitatives et un test t de Student a été utilisé pour les variables quantitatives. La survenue des décès de tous types a été analysée par la méthode de Kaplan-Meier.

Les résultats d'une analyse intermédiaire d'efficacité et de trois analyses intermédiaires de tolérance ont été revus par un comité indépendant de surveillance. Aucune modification de protocole n'a été recommandée par le comité au cours de l'étude.

RÉSULTATS

Populations de l'étude

Entre décembre 1996 et juillet 1998, 1 102 patients ont été inclus dans 60 centres de neuf pays. Au 14^e

jour, la phlébographie ou l'écho-Doppler de dépistage d'une thrombose veineuse profonde n'avaient pas été pratiqués chez 236 patients (Tableau 1). En conséquence, 866 patients ont été inclus dans l'analyse du critère de jugement principal, évalué à J14 par phlébographie chez 718 patients et par écho-Doppler chez 148 patients. Au 110^e jour, le critère d'évaluation secondaire avait été obtenu chez 798 patients (Tableau 1), dont 60,8 % directement et 39,2 % par téléphone. Parmi les 1 102 patients inclus dans l'étude, 1 073 avaient reçu au moins une dose du médicament à l'étude et avaient été inclus dans l'analyse de tolérance.

Caractéristiques des patients

Les caractéristiques initiales n'ont pas différé significativement entre le groupe placebo et les groupes énoxaparine (Tableau 2). Au total, 494 patients présentaient au moins deux raisons d'hospitalisation : 163 patients dans le groupe placebo, 159 dans le groupe énoxaparine 20 mg et 172 dans le groupe énoxaparine 40 mg. Globalement, 1 068 patients (96,9 % de la population de l'étude) avaient au moins un facteur de risque de maladie veineuse thrombo-embolique et le nombre moyen (\pm ET) de facteurs de risque par patient était de $2,1 \pm 1,1$ dans le groupe placebo, $2,0 \pm 1,1$ dans le groupe 20 mg et de $2,1 \pm 1,1$ dans le groupe 40 mg. La durée moyenne de traitement a été de 7 jours et n'a pas été significativement différente entre chacun des groupes énoxaparine et le groupe placebo.

TABLEAU 2. CARACTÉRISTIQUES INITIALES DES PATIENTS.*

CARACTÉRISTIQUES	PLACEBO (N = 371)	ÉNOXAPARINE 20 mg (N = 364)	ÉNOXAPARINE 40 mg (N = 367)
Âge — années	74,1 ± 10,6	72,9 ± 10,1	73,1 ± 10,8
Sexe — H/F	192/178	187/176	171/196
Indice de masse corporelle	25,0 ± 6,5	25,1 ± 6,2	24,9 ± 5,9
Motif de l'hospitalisation — nb. (%) †			
Insuffisance cardiaque congestive classe III NYHA	95 (25,7)	76 (20,9)	103 (28,1)
Insuffisance cardiaque congestive classe IV NYHA	32 (8,6)	44 (12,1)	26 (7,1)
Insuffisance respiratoire aiguë	202 (54,6)	192 (52,9)	195 (53,1)
Maladie infectieuse aiguë	193 (52,2)	194 (53,4)	197 (53,7)
Affection rhumatismale aiguë	32 (8,6)	40 (11,0)	28 (7,6)
Pathologie intestinale inflammatoire	1 (0,3)	1 (0,3)	3 (0,8)
Facteur de risque — nb. (%)			
Âge > 75 années	197 (53,2)	172 (47,4)	185 (50,4)
Cancer (antécédents ou actuel)	56 (15,1)	56 (15,4)	45 (12,3)
Antécédents thrombo-emboliques veineux	39 (10,5)	35 (9,6)	30 (8,2)
Obésité ‡	71 (19,2)	79 (21,8)	72 (19,6)
Varices	93 (25,1)	88 (24,2)	98 (26,7)
Traitement hormonal	9 (2,4)	8 (2,2)	5 (1,4)
Insuffisance cardiaque chronique	124 (33,5)	106 (29,2)	123 (33,5)
Insuffisance respiratoire chronique	197 (53,2)	197 (54,3)	195 (53,1)
≥ 2 Facteurs de risque — nb. (%)	247 (66,8)	241 (66,4)	245 (66,8)

* Les valeurs avec ± sont les moyennes ± l'écart-type. Les données de deux patients manquaient (un dans le groupe placebo et un dans le groupe 20 mg). Au total, 494 patients (respectivement, 163, 159 et 172 dans les groupes recevant le placebo, énoxaparine 20 mg et énoxaparine 40 mg) avaient deux raisons ou plus d'hospitalisation. NYHA signifie New York Heart Association.

† Les patients qui présentaient seulement une pathologie infectieuse aiguë, une arthrite aiguë ou une pathologie rhumatismale aiguë ou une maladie inflammatoire intestinale devaient avoir au moins un facteur de risque thrombo-embolique veineux additionnel pour être inclus dans l'étude.

‡ L'obésité était définie par un indice de masse corporelle d'au moins 30 pour les hommes et 28,6 pour les femmes.

Incidence des événements thrombo-emboliques veineux

L'incidence des événements thrombo-emboliques veineux au 14^e jour a été significativement plus faible dans le groupe énoxaparine 40 mg (5,5 % [16 patients sur 291]) que dans le groupe placebo (14,9 % [43 patients sur 288]) (risque relatif : 0,37 ; intervalle de confiance à 97,6 % : 0,22 à 0,63 ; $p < 0,001$) (Tableaux 3 et 4). Au 14^e jour, une embolie pulmonaire symptomatique non létale était survenue chez quatre patients, trois dans le groupe placebo et un dans le groupe 20 mg. Parmi les 100 thromboses veineuses profondes détectées au 14^e jour (dont 6 étaient symptomatiques), 92 ont été diagnostiquées à la phlébographie et 8 à l'écho-Doppler (3 symptomatiques et 5 asymptomatiques). L'incidence totale des thromboses veineuses profondes, celle des thromboses veineuses proximales et celle des distales ont été significativement plus faibles chez les patients du groupe 40 mg que chez ceux du groupe placebo. Il n'y a pas eu de différence significative concernant le critère principal d'évaluation entre les groupes 20 mg et placebo.

La réduction significative de l'incidence de l'ensemble des événements thrombo-emboliques veineux et des thromboses veineuses proximales et distales dans le groupe 40 mg s'est maintenue tout au long de la période des trois mois de suivi (Tableau 3). Huit événements

thrombo-emboliques veineux sont apparus entre les jours 15 et 110, dont quatre ont été des embolies pulmonaires mortelles : une dans le groupe placebo (trois semaines après l'interruption du traitement), une dans le groupe 20 mg et deux dans le groupe 40 mg (deux mois après l'interruption du traitement à l'étude).

Événements indésirables

Au 110^e jour, 142 patients étaient décédés : 50 dans le groupe placebo (13,9 %), 51 dans le groupe 20 mg (14,7 %) et 41 dans le groupe 40 mg (11,4 %) (Tableau 5). Le risque de mortalité a été plus faible dans le groupe 40 mg que dans le groupe placebo, mais cette différence n'a pas été significative (risque relatif : 0,83 ; intervalle de confiance à 95 % : 0,56 à 1,21 ; $p = 0,31$) (Figure 1). Il n'y a pas eu de différence significative sur le risque de mortalité entre les groupes 20 mg et placebo (risque relatif : 1,05 ; intervalle de confiance à 95 % : 0,71 à 1,56 ; $p = 0,80$). De plus, au 110^e jour, huit patients n'ayant pas reçu de traitement à l'étude étaient décédés (trois dans le groupe placebo et cinq dans le groupe 20 mg), mais les données sur ces patients n'ont pas été incluses dans l'analyse de tolérance.

Au cours de la période de traitement, une hémorragie majeure est survenue chez 11 patients (Tableau 5). Un patient du groupe 40 mg est décédé, mais l'hémor-

TABEAU 3. INCIDENCE DES ÉVÉNEMENTS THROMBO-EMBOLIQUES VEINEUX.

CRITÈRES	PLACEBO	ÉNOXAPARINE	
		20 mg	40 mg
		nombre (%)	
Critère principal			
Nb. patients évalués	288	287	291
Événements thrombo-emboliques veineux	43 (14,9)	43 (15,0)	16 (5,5)
Thrombose veineuse profonde seule	40 (13,9)	42 (14,6)	16 (5,5)
Embolie pulmonaire seule	2 (0,7)	0	0
Thrombose veineuse profonde et embolie pulmonaire	1 (0,3)	1 (0,3)	0
Thrombose veineuse profonde proximale	14 (4,9)	13 (4,5)	5 (1,7)
Thrombose veineuse profonde distale	27 (9,4)	30 (10,5)	11 (3,8)
Thrombose veineuse profonde symptomatique	2 (0,7)	3 (1,0)	1 (0,3)
Décès par embolie pulmonaire	0	0	0
Critère secondaire			
Nb. patients évalués	263	263	272
Événements thrombo-emboliques veineux	45 (17,1)	46 (17,5)	19 (7,0)
Thrombose veineuse profonde seule	41 (15,6)	44 (16,7)	17 (6,2)
Embolie pulmonaire seule	2 (0,8)	0	0
Thrombose veineuse profonde et embolie pulmonaire	1 (0,4)	1 (0,4)	0
Thrombose veineuse profonde proximale *	17 (6,5)	14 (5,3)	6 (2,2)
Thrombose veineuse profonde distale *	27 (10,3)	31 (11,8)	12 (4,4)
Thrombose veineuse profonde symptomatique	4 (1,5)	6 (2,3)	3 (1,1)
Décès par embolie pulmonaire	1 (0,4)	1 (0,4)	2 (0,7)

* En cas de thrombose veineuse profonde chez un même patient en cours de traitement et de nouveau au cours du suivi, la récurrence n'était pas comptabilisée comme un nouvel événement. Si le site (proximal ou distal) de la récurrence différait de celui du premier événement, les deux événements étaient indiqués séparément selon le site.

ragie n'a pas été considérée comme étant associée au traitement (elle était caractérisée par une hémoptysie massive liée à un carcinome bronchique). Il n'y a pas eu de cas d'hémorragie rétro-péritonéale ou intracrânienne au cours de la période de traitement. Au cours du suivi, deux patients supplémentaires sont décédés en raison d'une hémorragie, un dans le groupe 20 mg (d'une hématémèse massive) huit semaines après la fin du traitement et un dans le groupe 40 mg (d'hémorragie intracérébrale) trois semaines après la cessation du traitement.

Parmi les 31 cas de thrombopénie au cours de la période de traitement, 14 ont été considérés comme étant associés de façon possible ou probable au traitement ; ces 14 cas concernaient 8 patients dans le groupe placebo, 4 dans le groupe 20 mg et 2 dans le groupe 40 mg. Trois patients du groupe placebo ont présenté une thrombopénie sévère (Tableau 5). La thrombopénie a été associée à une thrombose veineuse ou artérielle chez cinq patients (quatre dans le groupe placebo et un dans le groupe 20 mg) au cours de la période de traitement. Il n'y a pas eu de différences significatives entre les deux groupes énoxaparine et le groupe placebo concernant l'incidence de survenue d'un autre événement intercurrent au cours de la période de suivi.

DISCUSSION

Dans notre étude contrôlée contre placebo, la survenue de la maladie thrombo-embolique veineuse chez des patients hospitalisés présentant une pathologie médicale aiguë a été significativement réduite par des injections quotidiennes d'énoxaparine 40 mg, sans augmentation du risque d'hémorragie majeure. Les incidences de 14,9 % des événements thrombo-emboliques veineux et de 4,9 % de thromboses veineuses profondes proximales dans le groupe placebo au cours de la période de traite-

TABEAU 4. RISQUES RELATIFS DE SURVENUE DES CRITÈRES PRINCIPAUX ET SECONDAIRES DE JUGEMENT POUR CHAQUE GROUPE D'ÉNOXAPARINE PAR RAPPORT AU GROUPE PLACEBO.*

CRITÈRES	ÉNOXAPARINE 20 mg		ÉNOXAPARINE 40 mg	
	RISQUE RELATIF IC 95 %	P	RISQUE RELATIF IC 95 %	P
Critère principal				
Événements thrombo-emboliques veineux ††	1,02 (0,70-1,51)	0,90	0,37 (0,22-0,63)	< 0,001
Thrombose veineuse profonde seule §	1,05 (0,71-1,57)	0,81	0,40 (0,23-0,69)	< 0,001
Thrombose veineuse profonde proximale §	0,93 (0,45-1,94)	1,00	0,35 (0,13-0,97)	0,04
Thrombose veineuse profonde distale §	1,11 (0,68-1,83)	0,68	0,40 (0,20-0,80)	0,01
Critère secondaire				
Événements thrombo-emboliques veineux ¶	1,02 (0,70-1,49)	0,91	0,41 (0,25-0,68)	< 0,001
Thrombose veineuse profonde seule §	1,07 (0,73-1,58)	0,81	0,40 (0,23-0,69)	< 0,001
Thrombose veineuse profonde proximale §	1,15 (0,71-1,88)	0,58	0,43 (0,22-0,84)	0,01
Thrombose veineuse profonde distale §	1,15 (0,71-1,88)	0,58	0,43 (0,22-0,84)	0,01

* IC signifie intervalle de confiance.

† Dans l'analyse du critère principal par rapport à l'ensemble des événements thrombo-emboliques veineux, les extrêmes entre parenthèses correspondent à l'intervalle de confiance à 97,6 % (voir chapitre Méthodes).

‡ Le test du chi-2 de Mantel-Haenszel a été utilisé.

§ Le test exact de Fisher a été utilisé.

¶ Le test du chi-2 a été utilisé.

TABEAU 5. INCIDENCE DES ÉVÉNEMENTS INDÉSIRABLES.*

ÉVÉNEMENTS INDÉSIRABLES	PLACEBO	ÉNOXAPARINE	
		20 mg	40 mg
		nombre (%)	
Période de traitement (J1-J14)			
Nb. patients évalués	362	351	360
Décès de toutes causes	16 (4,4)	15 (4,3)	12 (3,3)
Hémorragie †	31 (8,6)	41 (11,7)	45 (12,6)
Majeure	4 (1,1)	1 (0,3)	6 (1,7)
Fatale	0	0	1 (0,3)
Mineure	27 (7,5)	40 (11,4)	39 (10,8)
Réaction locale au site d'injection (hématome > 5 cm de diamètre)	0	4 (1,1)	5 (1,4)‡
Thrombopénie §	13 (3,6)	10 (2,8)	8 (2,2)
Thrombopénie sévère ¶	3 (0,8)	0	0
Période d'étude (J1-J110)			
Nb. patients évalués	362	351	360
Décès de toute cause	50 (13,9)	51 (14,7)	41 (11,4)
Hémorragie **	51 (14,3)	59 (17,2)	62 (17,4)
Majeure	7 (2,0)	4 (1,2)	12 (3,4)
Fatale	0	1 (0,3)	2 (0,6)
Mineure	45 (12,6)	57 (16,6)	51 (14,4)
Thrombopénie ††	17 (4,8)	11 (3,2)	10 (2,8)
Thrombopénie sévère ¶	3 (0,8)	0	0

* Les données de huit patients n'ayant reçu aucun médicament de l'étude et décédés au 110^e jour (trois patients dans le groupe placebo et cinq dans le groupe 20 mg) n'ont pas été incluses dans l'analyse des événements indésirables.

† Si les patients présentaient des hémorragies à la fois majeures et mineures, ils étaient comptés séparément dans ces catégories mais une seule fois dans les hémorragies totales. Des données sur les hémorragies au cours de la période de traitement sont manquantes chez trois patients (un dans le groupe placebo et deux dans le groupe 40 mg).

‡ P = 0,03 pour la comparaison avec le groupe placebo.

§ Des données sur les thrombopénies au cours de la période de traitement sont manquantes chez quatre patients (un dans le groupe placebo et trois dans le groupe 40 mg).

¶ Une thrombopénie sévère était définie par un nombre de plaquettes inférieur à 50 000/mm³.

|| Des données sur les décès de toutes causes au cours de la période d'étude sont manquantes chez six patients (un dans le groupe placebo, quatre dans le groupe 20 mg et un dans le groupe 40 mg).

** Si les patients présentaient des hémorragies à la fois majeures et mineures, ils étaient comptés séparément dans ces catégories mais une seule fois dans les hémorragies totales. Des données sur les hémorragies au cours de la période d'étude sont manquantes chez 18 patients (6 dans le groupe placebo ; 8 dans le groupe 20 mg et un dans le groupe 40 mg).

†† Des données sur les thrombopénies au cours de toute la période d'étude sont manquantes chez 19 patients (6 dans le groupe placebo, 9 dans le groupe 20 mg et 4 dans le groupe 40 mg).

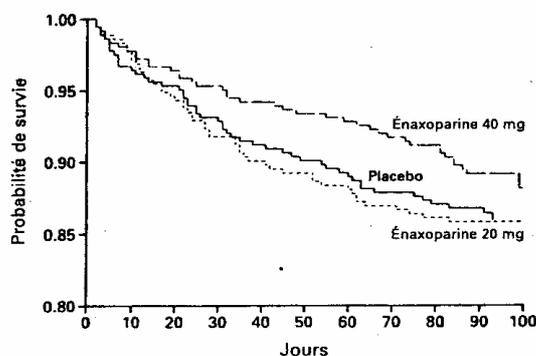
ment indiquent que cette population de patients présente un risque modéré de maladie veineuse thrombo-embolique selon la classification du Thromboembolism Risk Factors Consensus Group.¹³ Ces valeurs tombent dans l'écart de 9 à 26 % observé dans des études plus limitées.^{14,21,22,24} L'utilisation d'un groupe placebo a été considéré comme justifié sur le plan éthique car l'incidence de la maladie thrombo-embolique veineuse n'a pas été établie chez de tels patients, qu'aucune méthode de thrombo-prophylaxie chez ces patients n'a été établie et parce que les patients présentant un risque thrombo-embolique élevé avaient été exclus.^{3,4,13} Bien que de

faibles doses d'héparine non fractionnée soient largement utilisées dans la prophylaxie des thromboses, elles ne pouvaient être considérées comme un traitement contrôlé validé chez des patients hospitalisés en médecine. En effet, les quelques études soutenant l'utilisation de l'héparine non fractionnée ont inclus des petits nombres de patients,^{14,22,24} les résultats de deux études évaluant la mortalité chez des patients médicaux recevant 5 000 U deux fois par jour d'héparine non fractionnée sont controversés,^{18,20} et les recommandations des conférences de consensus ne sont pas définitives.^{13,27,32}

Notre observation d'une diminution de 63 % du risque d'événements thrombo-emboliques veineux dans le groupe recevant quotidiennement 40 mg d'énoxaparine est comparable à la diminution de 66 à 80 % rapportée dans des études plus limitées chez des patients médicaux recevant 5 000 U deux fois^{14,24} ou trois fois par jour²² de l'héparine non fractionnée ou énoxaparine 60 mg une fois par jour.²¹ Elle est également identique à la réduction habituellement observée lors de la prophylaxie avec l'héparine chez des patients chirurgicaux.^{3,4} Dans notre étude, l'incidence des événements thrombo-emboliques veineux au cours du traitement dans le groupe 40 mg a été inférieure à celle rapportée dans les deux groupes d'une étude récente, comparant 40 mg d'énoxaparine une fois par jour et 5 000 U d'héparine non fractionnée trois fois par jour chez des patients présentant une insuffisance respiratoire sévère ou une insuffisance cardiaque classe III ou IV selon la classification de la New York Heart Association ;³³ dans cette étude, l'incidence de la maladie thrombo-embolique veineuse était de 8,4 % dans le groupe énoxaparine et de 10,4 % dans le groupe héparine non fractionnée. La différence de taux peut s'expliquer par le fait que les patients dans cette étude étaient plus sévères que dans la nôtre. Nous n'avons pas observé de diminution de l'incidence de maladie thrombo-embolique veineuse avec la dose de 20 mg d'énoxaparine, bien que l'efficacité de cette dose soit identique à celle d'une dose de 5 000 U deux fois par jour d'héparine non fractionnée dans une autre étude concernant des patients médicaux probablement moins sévèrement affectés que dans notre étude ;²³ cette dose est également connue pour être efficace chez des patients chirurgicaux présentant un risque modéré.^{34,36}

Dans notre étude, le diagnostic de thrombose veineuse profonde a été principalement porté par phlébographie, qui demeure la méthode de dépistage de référence des thromboses veineuses profondes chez les patients asymptomatiques. Les ultrasons ont une faible sensibilité chez les patients asymptomatiques sous traitement médical^{15,16} ou en chirurgie orthopédique.³⁷ Vingt-deux pour cent des patients n'ont pas été inclus dans l'analyse du critère principal de jugement pour un certain nombre de raisons, la plupart étant liées à la sévérité relative de la pathologie et à la difficulté clinique de réalisation des phlébographies chez ces patients. Cependant, le nombre de patients qui n'ont pas été pris en compte pour l'ana-

ÉNOXAPARINE VS. PLACEBO DANS LA PRÉVENTION DE LA MALADIE THROMBO-EMBOLIQUE VEINEUSE CHEZ DES PATIENTS MÉDICAUX AIGUS



Nb. À RISQUE						
Total	1073	1022	983	965	943	231
Placebo	362	344	329	322	314	77
Énoxaparine 20 mg	351	332	316	310	302	81
Énoxaparine 40 mg	360	346	338	333	327	73

Figure 1. Courbes de Kaplan – Meier sur la probabilité de survie.
 Le risque de décès était plus faible dans le groupe énoxaparine 40 mg que dans le groupe placebo
 (risque relatif : 0,83 ; intervalle de confiance à 95 % : 0,56 à 1,21 ; p = 0,31).

lyse du critère principal de jugement, a été similaire dans les trois groupes.

La pertinence clinique des thromboses veineuses profondes asymptomatiques, en particulier distales, documentées par des tests objectifs, a été mise en question. Comme attendu, les événements symptomatiques ont été rares dans notre étude. De plus, la plupart des patients diagnostiqués comme ayant une thrombose veineuse asymptomatique au 14^e jour ont reçu des anticoagulants à dose curative. L'évolution naturelle de la maladie a donc pu être modifiée et l'association entre la diminution de l'incidence des événements asymptomatiques et les événements cliniques n'a pu être évaluée. Bien que trois études aient évalué la capacité de la thrombo-prophylaxie à réduire le risque de décès parmi les patients hospitalisés en médecine,^{18,20} les résultats sont encore controversés en raison des méthodologies utilisées. Notre étude, qui n'était pas conçue pour évaluer une différence dans la mortalité, a révélé une tendance clinique justifiée, avec une réduction absolue de 2,5 % du risque global des décès à trois mois dans le groupe énoxaparine 40 mg.

Nous avons choisi une durée de prophylaxie de 6 à 14 jours de façon à avoir une correspondance avec la durée habituelle d'hospitalisation des patients en médecine. Nous ne pouvons éliminer la possibilité que le traitement ait été trop court chez certains patients et qu'il ait été interrompu alors qu'ils présentaient encore un risque de maladie thrombo-embolique veineuse. En effet, deux embolies pulmonaires mortelles sont survenues dans le groupe 40 mg après plusieurs semaines d'interruption de la prophylaxie. D'un point de vue

pratique, la décision de prolonger la prophylaxie devrait donc être faite individuellement. Les études futures pourraient examiner les effets de la prolongation de la prophylaxie chez ces patients comme cela a été fait chez les patients subissant une intervention de la hanche.³⁸⁻⁴⁰

Soutenu par une bourse de Rhône-Poulenc Rorer (France).

ANNEXE

Les investigateurs et centres suivants ont participé à l'étude : Comité Scientifique — M.M. Samama (président), A.T. Cohen, J.-Y. Darmon, L. Desjardins, A. Eldor, C. Janbon, C.-G. Olsson, A.G. Turpie, N. Weisslinger (directeur du projet) ; Comité de Rédaction — A.T. Cohen, A. Leizorovicz, M.M. Samama, A.G. Turpie, N. Weisslinger ; Comité de Surveillance — A. Leizorovicz (président), H. Decousus, T. Lecompte ; Comité des Événements Critiques — Y. Gruel (président), C. Lamer, F. Parent ; Comité de Relecture Centralisée — P. Girard (président), M.-A. Collignon, P. Lacombe, D. Musset, M. Wartski ; Gestion du Projet — H. Nguyen (directeur de projet adjoint), C. Dole, N. Esposito, L. Laperrière ; Centre de Traitement des Données et des Analyses Statistiques — Unité de pharmacologie clinique, Hôpital Neuro-Cardiologique, Lyons, France : F. Boutitie (directeur de projet et statisticien), V. Bost (médecin événements critiques), E. Gauthier (coordinateur de l'étude), A. Chérif (secrétaire), M. Hervé (data manager) ; Coordination du Monitoring — Rhône – Poulenc Rorer (France) : A. Bone, A. Dal Pra, F. Kogan, F. Le Barbenchon ; Chiltern International, London : E. Delisle, T.K. Sohal, N. Spinnewyn ; Investigateurs — Canada (246 patients, 17 centres) : M. Alexander, D. Anderson, P. Brill-Edwards, C. Demers, R. Delage, L. Desjardins, S. Desmarais, M. Fitzgerald, R. Abboud, K. Grewal, J. Kassis, A. Kirby, S. Martel, J. Muscedere (2 centres), D. Rolf, R. Anderton, A.G. Turpie, J. Weitz, P. Wells ; France (163 patients, 17 centres) : J.-F. Bergmann, G. Simonneau, I. Mahé, C. Bonnamour, P.-L. Caraman, A. Chapellier, P. Cherin, J.-D. De Korwin, M. Morrisser, P. Dellamonica, F. Vandenbos, R. Dhote, C. Ginsburg, E. Duhamel, M. Gayraud, C. Janbon, I. Quéré, J.-J. Leduc, P. Letellier, H. Levesque, P. Mathern, C. Series, S. Guez, G. Thübau, D. Wahl ; Hongrie (61 patients, 2 centres) : A. Sasdi, I. Vaci ; Israël (28 patients, 3 centres) :

A. Eldor, J. Schwartz, G. Lugassi, J. Senderowicz ; Italie (36 patients, 3 centres) ; W. Ageno, L. Steidel, P. Prandoni, P. Bagatella, G. Valentini ; Slovaquia (22 patients, 2 centres) ; M. Hajkova, A. Okrucka ; Suède (108 patients, 3 centres) ; P. Hammarlund, R. Linne, C.-G. Olsson, A. Larsson ; Suisse (41 patients, 4 centres) ; H. Bounameaux, M.-J. Miron, A. Gallino, G. Plebani, D. Hayoz, T. Moccetti, B. Gorgio ; Royaume Uni (397 patients, 9 centres) ; A.T. Cohen, D.J. Quinlan (2 centres), A.J. Cowley, R. Storey, A. Darowski, R. Edmondson, P. Kesteven, W. MacNee, M. Skwarski, J. Turner, M. Siddiqui, J. Winter.

BIBLIOGRAPHIE

1. Lindblad B, Sternby NH, Bergqvist D. Incidence of venous thromboembolism verified by necropsy over 30 years. *BMJ* 1991 ; 302 : 709-11.
2. Sandler DA, Martin JF. Autopsy proven pulmonary embolism in hospital patients : are we detecting enough deep vein thrombosis ? *J R Soc Med* 1989 ; 82 : 203-5.
3. Clagett GP, Anderson FA Jr, Geerts WH, et al. Prevention of venous thromboembolism. *Chest* 1998 ; 114 : Suppl : 531S-560S.
4. Prevention of venous thromboembolism : international consensus statement (guidelines according to scientific evidence). *Int Angiol* 1997 ; 16 : 3-38.
5. Leizorovicz A, Haugh MC, Chapuis FR, Samama MM, Boissel JP. Low molecular weight heparin in prevention of perioperative thrombosis. *BMJ* 1992 ; 305 : 913-20.
6. Collins R, Scrimgeour A, Yusuf S, Peto R. Reduction in fatal pulmonary embolism and venous thrombosis by perioperative administration of subcutaneous heparin : overview of results of randomized trials in general, orthopedic, and urologic surgery. *N Engl J Med* 1988 ; 318 : 1162-73.
7. Oster G, Tuden RL, Colditz GA. Prevention of venous thromboembolism after general surgery : cost-effectiveness analysis of alternative approaches to prophylaxis. *Am J Med* 1987 ; 82 : 889-99.
8. Idem. A cost-effectiveness analysis of prophylaxis against deep-vein thrombosis in major orthopedic surgery. *JAMA* 1987 ; 257 : 203-8.
9. Bergqvist D, Matsch T, Jendteg S, Lindgren B, Persson U. The cost-effectiveness of prevention of post-operative thromboembolism. *Acta Chir Scand Suppl* 1990 ; 556 : 36-41.
10. Sandercock PAG, van den Belt AGM, Lindley RJ, Slattery J. Antithrombotic therapy in acute ischaemic stroke : an overview of the completed randomised trials. *J Neurol Neurosurg Psychiatry* 1993 ; 56 : 17-25.
11. Anderson FA Jr, Wheeler HB, Goldberg RJ, Hosmer DW, Forcier A. The prevalence of risk factors for venous thromboembolism among hospital patients. *Arch Intern Med* 1992 ; 152 : 1660-4.
12. Samama MM, Simonneau G, Wainstein J-P, De Vathaire F, Huet Y, Landauer D. SIRIUS Study : epidemiology of risk factors of deep venous thrombosis (DVT) of the lower limbs, in community practice. *Thromb Haemost* 1993 ; 69 : 763. abstract.
13. Thromboembolic Risk Factors (THRIFT) Consensus Group. Risk of and prophylaxis for venous thromboembolism in hospital patients. *BMJ* 1992 ; 305 : 567-74.
14. Cade JF. High risk of the critically ill for venous thromboembolism. *Crit Care Med* 1982 ; 10 : 448-50.
15. Keller F, Flosbach CW, PRIME Study Group. A randomised multicentre study investigating the efficacy and safety of the low molecular weight heparin enoxaparin versus unfractionated heparin in the prevention of thromboembolism in immobilised medical patients. *Thromb Haemost* 1995 ; 73 : 1106. abstract.
16. Harenberg J, Roehrborn P, Heene DL. Subcutaneous low-molecular-weight heparin versus standard heparin and the prevention of thromboembolism in medical inpatients. *Haemostasis* 1996 ; 26 : 127-39.
17. Harenberg J, Kallenbach B, Martin U, et al. Randomized controlled study of heparin and low molecular weight heparin for prevention of deep-vein thrombosis in medical patients. *Thromb Res* 1990 ; 59 : 639-50.
18. Halkin H, Goldberg J, Modan M, Modan B. Reduction of mortality in general medical inpatients by low-dose heparin prophylaxis. *Ann Intern Med* 1982 ; 96 : 561-5.
19. Bergmann JF, Caulin C. Heparin prophylaxis in bedridden patients. *Lancet* 1996 ; 348 : 205-6.
20. Gårdlund B. Randomised, controlled trial of low-dose heparin for prevention of fatal pulmonary embolism in patients with infectious disease. *Lancet* 1996 ; 347 : 1357-61.
21. Dahan R, Houlbert D, Caulin C, et al. Prevention of deep vein thrombosis in elderly medical in-patients by a low molecular weight heparin : a randomized double-blind trial. *Haemostasis* 1986 ; 16 : 159-64.
22. Belch JJ, Lowe GDO, Ward AG, Forbes CD, Prentice CRM. Prevention of deep vein thrombosis in medical patients by low-dose heparin. *Scott Med J* 1981 ; 26 : 115-7.
23. Bergmann JF, Neuhart E. A multicenter randomized double-blind study of enoxaparin compared with unfractionated heparin in the prevention of venous thromboembolic disease in elderly in-patients bedridden for an acute medical illness. *Thromb Haemost* 1996 ; 76 : 529-34.
24. Ibarra-Pérez C, Lau-Cortes E, Colmenero-Zubiate S, et al. Prevalence and prevention of deep venous thrombosis of the lower extremities in high-risk pulmonary patients. *Angiology* 1988 ; 39 : 505-13.
25. The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism : results of the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED). *JAMA* 1990 ; 263 : 2753-9.
26. Statistical methods in epidemiology. In : Armitage P, Berry G. Statistical methods in medical research. 2nd ed. Oxford, England : Blackwell Scientific, 1987 : 455-83.
27. Hirsh J, Fuster V. Guide to anticoagulant therapy. 1. Heparin. *Circulation* 1994 ; 89 : 1449-68.
28. Goldhaber SZ, Morpurgo M. Diagnosis, treatment, and prevention of pulmonary embolism : report of the WHO/International Society and Federation of Cardiology Task Force. *JAMA* 1992 ; 268 : 1727-33.
29. Prevention of venous thrombosis and pulmonary embolism : NIH Consensus Development. *JAMA* 1986 ; 256 : 744-9.
30. Nicolaidis AN. Prevention of venous thromboembolism : European consensus statement, 15 November 1997, developed at Oakley Court Hotel, Windsor, UK. *Int Angiol* 1992 ; 11 : 151-9.
31. Clagett GP, Anderson FA Jr, Levine MN, Salzman EW, Wheeler HB. Prevention of venous thromboembolism. *Chest* 1992 ; 102 : Suppl : 391S-407S.
32. Clagett GP, Anderson FA Jr, Heit J, Levine MN, Wheeler HB. Prevention of venous thromboembolism. *Chest* 1995 ; 108 : Suppl : 312S-334S.
33. Kleber FX, Witt C, Flosbach CW, Koppenhagen K, Vogel G, PRINCE Study Group. Study to compare the efficacy and safety of the LMWH enoxaparin and standard heparin in the prevention of thromboembolic events in medical patients with cardiopulmonary diseases. *Ann Hematol* 1998 ; 76 : Suppl 1 : A93. abstract.
34. Samama MM, Bernard P, Bonnardot JP, Combe-Tanzali S, Lanson Y, Tissot E. Low molecular weight heparin compared with unfractionated heparin in prevention of postoperative thrombosis. *Br J Surg* 1988 ; 75 : 128-31.
35. Gazzaniga GM, Angelini G, Pastorino G, et al. Enoxaparin in the prevention of deep venous thrombosis after major surgery : multicentric study. *Int Surg* 1993 ; 78 : 271-5.
36. Nurmohamed MT, Verhaeghe R, Haas S, et al. A comparative trial of a low molecular weight heparin (enoxaparin) versus standard heparin for the prophylaxis of postoperative deep vein thrombosis in general surgery. *Am J Surg* 1995 ; 169 : 567-71.
37. Wells PS, Lensing AWA, Davidson BL, Prins MH, Hirsh J. Accuracy of ultrasound for the diagnosis of deep venous thrombosis in asymptomatic patients after orthopedic surgery : a meta-analysis. *Ann Intern Med* 1995 ; 122 : 47-53.
38. Bergqvist D, Benoni G, Björgell O, et al. Low-molecular-weight heparin (enoxaparin) as prophylaxis against venous thromboembolism after total hip replacement. *N Engl J Med* 1996 ; 335 : 696-700.
39. Planes A, Vochelle N, Darmon JY, Fagola M, Bellaud M, Huet Y. Risk of deep venous thrombosis after hospital discharge in patients having undergone total hip replacement : double-blind randomised comparison of enoxaparin versus placebo. *Lancet* 1996 ; 348 : 224-8.
40. Dahl OE, Andreassen G, Asplund T, et al. Prolonged thromboprophylaxis following hip replacement surgery : results of a double-blind, prospective, randomised, placebo-controlled study with dalteparin. *Thromb Haemost* 1997 ; 77 : 26-31.

Enoxaparine comparée à un placebo dans la prévention de la maladie thrombo-embolique veineuse chez des patients présentant une pathologie médicale aiguë.

Question 1 :

Identifiez les deux questions auxquelles cherchait à répondre ce travail ? Le type d'étude utilisée est-il optimum pour répondre à la question posée ?

Question 2 :

Quel est le résultat mis en avant par les auteurs comme apport de leur travail ?

Question 3 :

Expliquez, à partir des résultats rapportés dans le tableau III, pourquoi la dose 40 mg a été retenue à l'issue de l'essai ?

Question 4 :

Déterminez le nombre de patients randomisés qui ne participe pas au résultat obtenu sur le critère principal. Quel type de biais est susceptible d'induire ces patients ?

Question 5 :

Déterminez le nombre de patients randomisés qui ne participe pas au résultat obtenu sur le critère principal. Quel type de biais est susceptible d'induire ces patients ?

Question 6 :

Pouvez-vous utiliser ce résultat pour justifier l'utilisation de l'énoxaparine 40 mg pour la thromboprophylaxie chez un jeune homme de 30 ans immobilisé pour une fracture du membre inférieur sous plâtre ?

Question 7 :

Afin d'évaluer la pertinence clinique du résultat mis en avant, précisez quel est le type de critère de jugement utilisé (critère principal) et discutez sa pertinence clinique.