

FACULTE DE MEDECINE DE LIMOGES

---

# **LA LECTURE CRITIQUE D'ARTICLES (LCA)**

**PREPARATION A L'EXAMEN NATIONAL CLASSANT**

**DCEM2  
2009-2010**

Professeur A. VERGNENEGRE

Professeur P. M. PREUX

Docteur B. MARIN

L'épreuve de lecture critique d'articles (LCA) fait partie des épreuves de l'examen national classant.

Il s'agit d'une épreuve difficile, qui demande de la part des étudiants, une grande organisation :

- il faut d'abord respecter les trois heures de l'épreuve, c'est-à-dire s'organiser pour que la lecture de l'article, ainsi que la mise en évidence des passages importants ne pénalise pas l'étudiant dans la rédaction des réponses aux questions,

- il faut garder le temps nécessaire pour répondre aux questions posées qui comptent pour 80% de la note, mais également pour répondre au résumé qui compte pour 20% de la note.

Le but de ce document est de présenter les textes réglementaires qui organisent cette épreuve et également de leur donner les bases nécessaires et suffisantes pour pouvoir répondre aux questions et écrire le résumé.

**CHAPITRE I :**

**LES TEXTES REGLEMENTAIRES**

## ÉPREUVE DE LECTURE CRITIQUE D'UN ARTICLE MÉDICAL

L'objectif de l'épreuve est d'amener l'étudiant à lire de façon critique et à analyser le contenu d'un article en vue de son autoformation actuelle et future. Le mot critique ne doit pas être entendu dans le sens où l'on demanderait aux étudiants de chercher systématiquement tous les défauts d'un article. Cette épreuve part du principe que toute information médicale doit être analysée avec du recul, en cherchant les défauts éventuels mais aussi les limites, les implications, l'utilité pour la pratique.

L'article choisi doit être considéré comme de bonne qualité (le but de l'enseignement de la lecture critique est de donner à la fois le sens critique et le goût de la lecture).

### I – PRINCIPES DE L'ÉPREUVE

L'épreuve porte sur des articles :

- ❖ Médicaux scientifiques originaux, c'est-à-dire rapportant une étude d'observation ou expérimentale, traitant d'un sujet mentionné dans le programme de la deuxième partie du 2<sup>ème</sup> cycle (bulletin officiel du Ministère de l'Éducation Nationale n° 31 du 30 août 2001 à l'adresse [www.education.gouv.fr/bo/2001/31/sup.htm](http://www.education.gouv.fr/bo/2001/31/sup.htm))
- ❖ Issus d'une revue médicale avec comité de lecture,
- ❖ En français (articles primitivement en français ou traduits d'une autre langue).
- ❖ Utilisés en entier (ce qui est souhaitable) ou partiellement s'ils sont trop longs.
- ❖ Seront supprimés :
  - Le résumé de l'article puisqu'élaborer un résumé est constitutif de l'épreuve ;
  - Le noms des auteurs et la référence de la revue dans la mesure où l'article peut être traduit ou constituer un extrait du texte original.

### II – DEROULEMENT DE L'ÉPREUVE

Rédactionnelle, elle dure 3 heures et compte pour 10% de la note totale des épreuves classantes nationales (5% en 2009 puis 10% dès 2010). Les réponses se font avec des phrases construites ; le style télégraphique n'est pas accepté. Les abréviations ne sont pas admises sauf celles qui figurent sur le site du CNCI.

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

**L'épreuve comporte deux parties** : un résumé (20% de la note) et des questions (80 % de la note).

1°) Le résumé doit être structuré en fonction du contenu de l'article. Il comporte généralement : objectifs de l'étude, matériel et méthodes, résultats et conclusion. **Il ne doit pas dépasser 250 mots.**

2°) Les questions sont au nombre de 6 à 10 et portent sur les objectifs pédagogiques ci-dessous. La longueur des réponses peut varier selon le type de question et n'est pas limitée par un cadre : en moyenne, la question doit amener l'étudiant à répondre sous forme de phrases.

### III – CONSEILS POUR L'ELABORATION DE L'EPREUVE

#### A – OBJECTIFS PEDAGOGIQUES

##### - IDENTIFIER

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 à l'exclusion des meta-analyses.

2°) La « question » posée par les auteurs (hypothèse)

##### - ANALYSER LA METHODOLOGIE

###### ❖ Population étudiée :

3°) Identifier les caractéristiques (données démographiques) de la population étudiée, à laquelle, les conclusions pourront être appliquées.

4°) Analyser les modalités de sélection des sujets, critères d'inclusion et de non inclusion et d'exclusion.

5°) Identifier la technique de randomisation et vérifier sa cohérence, le cas échéant.

6°) Discuter la comparabilité des groupes soumis à la comparaison.

7°) Discuter l'évolution des effectifs étudiés et leur cohérence dans la totalité de l'article ; savoir si le calcul du nombre de sujets nécessaires a été effectué à priori

###### ❖ Méthodes

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.

9°) Vérifier que les analyses statistiques (en fonction de notions élémentaires) sont cohérentes avec le projet du travail ; connaître les limites de l'analyse par sous groupe ; connaître la notion de perdus de vue.

10°) Vérifier le respect des règles d'éthique.

##### - ANALYSER LA PRESENTATION DES RESULTATS

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

12°) Vérifier la présence des indices de dispersion permettant d'évaluer la variabilité des mesures et de leurs estimateurs.

##### - CRITIQUER L'ANALYSE DES RESULTATS ET DE LA DISCUSSION

13°) Discuter la nature et la précision des critères de jugement des résultats.

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

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

16°) Discuter la signification statistique des résultats.

17°) Discuter la pertinence clinique des résultats.

18°) Vérifier que les résultats offrent une réponse à la question annoncée.

19°) Vérifier que les conclusions sont justifiées par les résultats.

20°) Indiquer le niveau de preuve de l'étude (grille de l'HAS).

**- EVALUER LES APPLICATIONS CLINIQUES**

21°) Discuter la ou les applications potentielles proposées par l'étude

**- ANALYSER LA FORME DE L'ARTICLE**

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.

23°) Faire une analyse critique de la présentation des références.

24°) Faire une analyse critique du titre.

**B – 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.
- ❖ Les questions ne sont pas posées dans un langage de spécialistes. Elles doivent être comprises simplement par tout étudiant de deuxième cycle. Elles sont indépendantes les unes des autres.
- ❖ Il faut éviter les sous questions.
- ❖ 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.

**IV – MODALITES DE CORRECTION**

- ❖ Le résumé ne doit pas dépasser 250 mots et doit être fidèle à l'article. Il sera proposé un cadre matérialisé. Une pénalité sera décidée, le cas échéant, par le jury en cas de dépassement, proportionnelle au degré de dépassement.
- ❖ Une grille de réponses est proposée par le conseil scientifique du CNCI pour le résumé et les questions portant sur l'article
- ❖ Des points négatifs peuvent être décidés par le jury.
- ❖ L'épreuve fait l'objet d'une double correction comme pour les dossiers.

Le site du CNCI.

<http://www.cnci.univ-paris5.fr/medecine/>

Sur le site sont disponibles :

- Grille pour compter les mots (nombre de mots dans la colonne de gauche)

<http://www.cnci.univ-paris5.fr/medecine/Resumer-LCA.pdf>

- Définition de ce qu'est un mot, ce qui peut être mis dans une seule case

<http://www.cnci.univ-paris5.fr/medecine/Compter-LCA.pdf>

- Liste des abréviations utilisables d'emblée sans préalable définition dans le texte :

<http://www.cnci.univ-paris5.fr/medecine/abreviations.html>

**CHAPITRE II :**

**GENERALITES SUR LA LECTURE CRITIQUE D'UN ARTICLE  
MEDICAL**



**OBJECTIFS:**

Il s'agit de la lecture d'un article scientifique original, de bonne qualité et en français.

Cet article concerne l'un des domaines cliniques du programme de l'ENC.

Les objectifs pédagogiques de l'épreuve sont :

- la bonne compréhension de l'intérêt d'une publication scientifique pour la connaissance médicale,

- la déduction éventuelle d'une application de ces connaissances nouvelles pour la pratique.

En aucun cas, il ne s'agit de transformer l'étudiant du 2<sup>ème</sup> cycle en expert.

Il s'agit d'une lecture raisonnée et raisonnable.

Le déroulement de l'épreuve est la suivante : épreuve de trois heures, rédactionnelle, avec deux parties : écriture d'un résumé de moins de 250 mots qui compte pour 20% de la note et habituellement entre six à huit questions qui comptent pour 80% de la note.

**LES FONDAMENTAUX :**

- identifier l'objet, l'hypothèse et la question de l'étude : étude épidémiologique de prévalence ou d'incidence, d'un facteur de risque, nouvelle approche d'une stratégie diagnostique ou d'une stratégie thérapeutique, programme de prévention, dépistage, éducation, facteurs pronostics d'une maladie, ...

- analyser la méthodologie de l'article avec les critères d'inclusion, d'exclusion, les caractères cibles de la population, la validité de la randomisation, la comparabilité des groupes, le choix des effectifs

- analyser la cohérence entre les objectifs et le type de l'étude, entre les objectifs et les résultats, ainsi qu'avec les analyses statistiques réalisées

- analyser la présentation dans sa précision, clarté et cohérence

- analyser les résultats et la discussion : nature et précision des critères de jugement, biais identifiés pris en compte, signification statistique des résultats, conclusion ou recommandations soutenues par les résultats

- être capable d'indiquer le niveau de preuve de l'étude : cf tableau extrait du guide de lecture des articles médicaux de la Haute Autorité de Santé (HAS).

Niveau de preuve scientifique fourni par la littérature (études thérapeutiques)	Grade des recommandations
Niveau 1 (NP1) <ul style="list-style-type: none"> <li>• Essais comparatifs randomisés de forte puissance.</li> <li>• Méta-analyse d'essais comparatifs randomisés</li> <li>• Analyse de décision basée sur des études bien menées</li> </ul>	Preuve scientifique établie A
Niveau 2 (NP2) <ul style="list-style-type: none"> <li>• Essais comparatifs randomisés de faible puissance.</li> <li>• Études comparatives non randomisées bien menées</li> <li>• Études de cohorte</li> </ul>	Présomption scientifique B
Niveau 3 (NP3) <ul style="list-style-type: none"> <li>• Études Cas-témoin</li> </ul>	Faible niveau de preuve C
Niveau 4 (NP4) <ul style="list-style-type: none"> <li>• Études comparatives comportant des biais importants</li> <li>• Études rétrospectives</li> <li>• Séries de cas</li> </ul>	

Pour la rédaction des questions, aucun calcul nouveau ne doit être demandé. L'étudiant doit être capable d'analyser la forme de l'article (introduction, matériels, méthode, résultats et discussion). Un commentaire des tableaux peut être demandé. Les questions doivent être posées de façon simple, sans ambiguïté.

## **CHAPITRE III :**

# **L'ANALYSE GENERALE DES DIFFERENTS TYPES D'ARTICLES SCIENTIFIQUES ISSUS DE LA LITTERATURE**

**D'après (1)**

1) ANAES. Analyse de la littérature et gradation des recommandations. Agence Nationale d'Accréditation et d'Evaluation en Santé, 2000 ;1vol :64 p

Quelque soit l'article soumis à l'étudiant, la méthodologie d'analyse reste toujours la même. Un certain nombre de questions préliminaires doivent être posées.

## **I- LES QUESTIONS SPECIFIQUES A TOUS LES TYPES D'ARTICLES :**

⇒ *Pourquoi l'étude a-t-elle été entreprise ?*

Objectif et hypothèse de travail

⇒ *Quel est le type de l'étude ?*

- Etude expérimentale (animaux, volontaires sains, sujets malades), de type essai clinique (l'essai clinique est le plus souvent thérapeutique et il peut également s'intéresser aux dispositifs médicaux ou plus largement aux « stratégies thérapeutiques ou diagnostiques ». Souvent l'essai est comparatif (dit de supériorité) mais une équivalence entre deux traitements peu aussi être évaluée.
- Etude observationnelle (touchant une population ou un échantillon), pouvant être descriptive (évaluation d'une prévalence) ou analytique : étude de cohorte (étude pronostique) ou étude cas-témoin.
- Etude diagnostique pour évaluer les qualités d'un nouveau procédé diagnostique par rapport au glod standard par exemple.
- revue de la littérature : synthèse, méta-analyse,
- recommandations de pratique,
- analyse décisionnelle et économique

⇒ *Le protocole était-il approprié ?*

Pour le savoir, il faut se pencher sur la section « méthodes ». C'est la raison pour laquelle celle-ci doit être lue après le titre et le résumé. On essaiera de savoir :

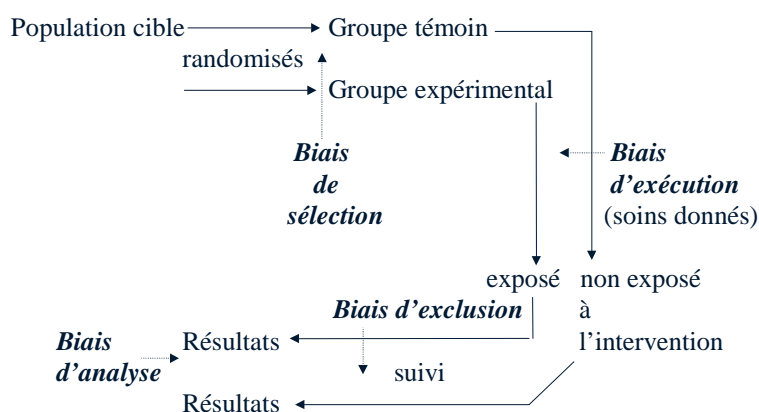
- s'il s'agit d'une étude originale, et notamment qu'elle est alors son apport par rapport aux travaux antérieurs.
- qui sont les sujets de l'étude (recrutement, inclusion, non-inclusion, perdu de vue).

⇒ *Le protocole a-t-il été judicieusement conçu ?*

- Stratégie étudiée,
- Critères de jugement.

⇒ *Comment les biais ont-ils été évités ou minimisés ?*

Ils doivent être décrits et commentés par les auteurs. Les biais sont décrits plus en détail dans le chapitre V.



⇒ *Les mesures ont-elles été réalisées en aveugle ?*

⇒ *La validité statistique est-elle assurée ?*

- taille de l'échantillon
- ampleur de la différence
- durée et taille du suivi

## **II- LES QUESTIONS SPECIFIQUES A CERTAINS TYPES D'ARTICLES :**

*A) Analyse d'un article thérapeutique :*

La méthodologie des essais thérapeutiques est traitée dans un autre polycopié.

La grille de lecture est présentée en annexe. Elle concerne essentiellement la méthodologie sur la nature de l'essai comparatif, prospectif, randomisé, sur les populations, les variables cliniquement pertinentes prises en compte, et surtout une analyse statistique adaptée (cf chapitre V). Une des questions importante est de savoir si l'analyse est faite en intention de traiter, c'est une analyse qui prend en compte tous les patients inclus dans l'étude, même si ceux-ci n'ont pas participé à la totalité du protocole (erreur, arrêt de traitement).

L'applicabilité clinique des résultats doit être discutée en fonction des significations et des modalités du traitement.

<b>GRILLE DE LECTURE D'UN ARTICLE THÉRAPEUTIQUE</b>
---

**Titre et auteur de l'article:** \_\_\_\_\_

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

	<b>OUI</b>	<b>NON</b>	<b>?</b>
1. Les objectifs sont clairement définis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Méthodologie de l'étude			
• L'étude est comparative	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- <i>l'étude est prospective</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- <i>l'étude est randomisée</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Le calcul du nombre de patients a été fait a priori	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• La population de l'étude correspond à la population habituellement traitée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Toutes les variables cliniquement pertinentes sont prises en compte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• L'analyse statistique est adaptée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• L'analyse est faite en intention de traiter	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Les résultats sont cohérents avec l'objectif de l'étude et tiennent compte d'éventuels effets secondaires	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Applicabilité clinique			
• La signification clinique est donnée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les modalités de traitement sont applicables en routine	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Commentaires :**

*B) Analyse d'un article diagnostic :*

Les critères recherchés pour apprécier la validité du diagnostic d'un test doivent apprécier : la capacité de ce test à mesurer effectivement ce que l'on veut mesurer et la variabilité des résultats liés à l'événement étudié.

Les objectifs doivent être définis ainsi que la méthodologie. L'analyse des résultats doit décrire la sensibilité, la spécificité, l'efficacité diagnostique qui correspond au pourcentage de biens classés par le test (elle est égale aux vrais positifs + les vrais négatifs / malades + non malades

$$= \frac{a + d}{a + b + c + d}$$

Les rapports de vraisemblance doivent éventuellement être utilisés pour quantifier la vraisemblance du diagnostic fournit par le test positif. Ce rapport de vraisemblance LR (Likelihood Ratio) correspond au rapport de la probabilité qu'un test positif corresponde réellement à une maladie par rapport à la probabilité qu'un test positif ne corresponde pas à la présence de cette maladie.

$$\text{LR d'un test positif} = \frac{\text{vrai positif/malades}}{\text{faux positifs/non malades}} = \text{sensibilité}/1 - \text{spécificité}.$$

$$\text{LR d'un test négatif} = \frac{\text{faux négatif/malades}}{\text{vrai négatifs/non malades}} = 1 - \text{sensibilité}/\text{spécificité}.$$

L'applicabilité du test hors tout contexte expérimental et son utilité doivent être décrites.

<b>GRILLE DE LECTURE D'UN ARTICLE DIAGNOSTIQUE</b>
--

**Titre et auteur de l'article:** \_\_\_\_\_

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

	OUI	NON	?
1. Les objectifs sont clairement définis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Méthodologie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Le test étudié est comparé à un test de référence fiable et valide, déterminé à priori	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• La méthode de sélection des patients est décrite	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• La fréquence de la maladie dans l'échantillon étudié correspond aux données épidémiologiques connues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Le terme « normal » est défini	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Analyse des résultats			
• Les résultats sont analysés en aveugle quand c'est possible	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les caractéristiques diagnostiques du test sont calculées ou calculables (sensibilité, spécificité)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. L'utilité clinique du test est recherchée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Commentaires :**



C) *Analyse d'un article de causalité (épidémiologie analytique) :*

La grille de lecture d'un article de causalité permet de répondre à des grandes questions. L'association existe-t-elle entre un facteur de risque et la maladie ?, cette relation est-elle forte ?, le facteur de risque peut-il être considéré comme causal ?.

Le schéma de l'étude doit être précisé, ainsi que sa méthodologie (l'étude est-elle comparative ?, les populations prises en compte sont-elles bien définies ?, les facteurs de risque et d'exposition sont-ils bien définis ?).

Les résultats doivent permettre de préciser, soit le risque relatif, soit l'odds-ratio en fonction du type d'étude. La causalité de l'association doit être discutée. Plusieurs aspects doivent être analysés :

- par rapport à l'étude, l'association est retrouvée de manière uniforme et logique dans l'analyse globale,
- une relation dose-effet peut être mise en évidence par rapport aux connaissances générales,
- l'association retrouvée dans l'étude est vraisemblable et en accord avec les connaissances épidémiologiques,
- les biais doivent être décrits et pris en compte dans la grille de lecture décrite ci-après.

<b>GRILLE DE LECTURE D'UN ARTICLE DE CAUSALITÉ (EPIDEMIOLOGIE ANALYTIQUE)</b>
---

**Titre et auteur de l'article:** \_\_\_\_\_

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

	OUI	NON	?
1. La formulation des objectifs est clairement exprimée			
2. Méthodologie			
• L'étude est comparative	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les populations exposées et témoins prises en compte sont bien définies (caractéristiques, critères d'inclusion et d'exclusion)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les facteurs de risques et d'exposition sont bien définis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Résultats			
• Les groupes sont comparés	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• L'existence d'une association est prouvée et la force de l'association est testée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• La causalité de l'association est étudiée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les biais sont décrits et pris en compte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Commentaires :**

D) *Analyse d'un article épidémiologique de cohorte :*

Les objectifs de l'étude doivent être clairement définis, ainsi que la méthodologie. Les modalités de constitution de la cohorte doivent être précisés. Les critères d'inclusion et d'exclusion doivent être spécifiés et adéquats.

Les biais possibles sont-ils pris en compte ?, le suivi est-il compris et correctement réalisé ?

La grille de lecture est décrite ci-après.

<b>GRILLE DE LECTURE D'UN ARTICLE EPIDEMIOLOGIQUE DE COHORTE</b>
--

**Titre et auteur de l'article:** \_\_\_\_\_

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

	OUI	NON	?
1. Les objectifs de l'étude sont clairement définis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Méthodologie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les modalités de constitution de la cohorte sont précisées.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Tous les patients de la cohorte ont été identifiés au même stade de la maladie.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les critères d'inclusion et d'exclusion sont spécifiés et adéquats.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les biais possibles sont exposés et les méthodes pour les prendre en compte sont décrites.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Le suivi est complet et correctement réalisé.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les critères de jugement sont pertinents, fiables et tous utilisés.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Les résultats			
• L'interprétation de ces critères est objective.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Les résultats sont ajustés pour les autres facteurs pronostiques.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Commentaires :**

**EN CONCLUSION :**

Lire un article scientifique repose donc sur une méthodologie relativement complexe. Le lecteur doit s'adapter au type d'article auquel il est confronté. La rigueur des grilles de lecture est importante si l'on veut, à partir d'une masse d'articles existants, pouvoir sélectionner les plus intéressants. Ceci évite la dispersion et fait gagner un gain de temps précieux au médecin. Ces grilles de lecture améliorent également un tri qui pourrait être subjectif.

**CHAPITRE IV :**

**LA REDACTION DU RESUME DANS L'EPREUVE DE LECTURE  
CRITIQUE D'ARTICLE**

Le résumé d'un article scientifique est important car il est lu juste après le titre. Il s'agit probablement de la partie qui est la plus lue de la littérature qui est produite.

Le rôle du résumé est d'informer de manière abrégée sur le contenu de l'article.

## I- LES DIFFERENTS TYPES DE RESUME :

On en distingue trois types qui sont reproduits sur le tableau ci-après :

**Tableau VII. Résumé d'une étude, présenté sous forme indicative, informative ou structurée.**

<i>Forme</i>	<i>Exemple*</i>
Indicative	Nous rapportons les résultats d'une étude dont l'objectif était de rechercher si la tuberculose joue un rôle dans la progression de l'infection par le virus de l'immunodéficience humaine. Les patients étudiés étaient issus de la Cohorte Aquitaine. Dans cet article, nous rapportons les résultats des analyses comparatives de la survie chez des patients tuberculeux et des sujets témoins. Nous discutons les implications pour la prise en charge thérapeutique et prophylactique de la tuberculose dans le cadre de l'infection par le virus de l'immunodéficience humaine.
Informative	Pour évaluer le rôle de la tuberculose dans l'infection par le virus de l'immunodéficience humaine (VIH), nous avons comparé la survie de patients tuberculeux et de sujets témoins appariés en fonction de l'année du diagnostic de l'infection par le VIH et le taux de lymphocytes CD4 <sup>+</sup> . Les patients étaient issus de la Cohorte Aquitaine, qui suit depuis 1985 tous les patients infectés par le VIH âgés de plus de 13 ans et vus dans cinq hôpitaux. L'inclusion des témoins se faisait à la date du diagnostic de la tuberculose du patient correspondant. Les patients qui avaient reçu une prophylaxie primaire contre les mycobactéries autres que la tuberculose ont été exclus. Au 30 juin 1994, 104 patients tuberculeux et 620 sujets témoins ont été sélectionnés ; ils étaient comparables sauf pour les antécédents de toxicomanie intraveineuse (patients tuberculeux : 51 % versus sujets témoins : 31 %) et d'infection opportuniste (40 % versus 29 %). La survie était plus courte chez les sujets tuberculeux que chez les sujets témoins (rapport de risque : 1,5 ; intervalle de confiance à 95 % : 1,2 ; 2,1) même après ajustement sur les différences à l'entrée. Le risque d'infection opportuniste ou de passage sous 50 lymphocytes CD4 <sup>+</sup> /mm <sup>3</sup> était légèrement plus élevé chez les patients tuberculeux que chez les sujets témoins, mais pas significativement. La tuberculose accélère peut-être l'évolution de l'infection par le VIH.
Structurée	<p><b>Objectif</b> - Évaluer le rôle de la tuberculose dans l'infection par le virus de l'immunodéficience humaine (VIH).</p> <p><b>Schéma d'étude</b> - Étude de cohorte.</p> <p><b>Population</b> - Patients tuberculeux et sujets témoins issus de la Cohorte Aquitaine, qui suit depuis 1985 tous les patients infectés par le VIH âgés de plus de 13 ans et vus dans cinq hôpitaux. Témoins appariés en fonction de l'année du diagnostic de l'infection par le VIH et le taux de lymphocytes CD4<sup>+</sup>, inclus à la date du diagnostic de la tuberculose du patient correspondant. Exclusion des patients ayant reçu une prophylaxie primaire contre les mycobactéries autres que la tuberculose.</p> <p><b>Effets mesurés</b> - Survie, survenue d'infection opportuniste, passage sous 50 CD4<sup>+</sup>/mm<sup>3</sup>.</p> <p><b>Résultats</b> - Au 30 juin 1994, 104 patients tuberculeux et 620 sujets témoins ont été sélectionnés ; ils étaient comparables sauf pour les antécédents de toxicomanie intraveineuse (patients tuberculeux : 51 % versus sujets témoins : 31 %) et d'infection opportuniste (40 % versus 29 %). La survie était plus courte chez les patients tuberculeux que chez les sujets témoins (rapport de risque : 1,5 ; intervalle de confiance à 95 % : 1,2 ; 2,1), même après ajustement sur les différences à l'entrée. Le risque d'infection opportuniste ou de passage sous 50 lymphocytes CD4<sup>+</sup>/mm<sup>3</sup> était légèrement plus élevé chez les patients tuberculeux que chez les sujets témoins, mais pas significativement.</p> <p><b>Conclusion</b> - La tuberculose accélère peut-être l'évolution de l'infection par le VIH.</p>

1) Une forme indicative, peu utilisée, qui présente un intérêt scientifique et médical extrêmement limité.

2) Une forme informative, qui a comme objectif de renseigner le lecteur sur le contenu de l'article. On distingue deux types de formes informatives :

- la forme informative non structurée qui est reproduite à la partie centrale du tableau ci-dessus,
- de nombreuses revues maintenant préfèrent une forme structurée où le rédacteur va présenter, par toute une série de mots clés, au lecteur, les informations importantes. Certaines revues imposent une structure alors résumée. Lorsque ceci n'est pas le cas, on peut séparer le résumé en utilisant quatre mots clés :
  - ✓ contexte (ou objectifs),
  - ✓ méthodes,
  - ✓ résultats,
  - ✓ conclusion.



## **II- EPREUVE DE RESUME DE LCA A L'ECN :**

Ce résumé doit comporter 250 mots maximum. Pour cela, une grille est à la disposition de l'étudiant sur laquelle il doit écrire un mot par case. Ce résumé sera structuré. Sa non structuration entraîne une sanction au niveau des points attribués.

Le résumé doit être fidèle à l'article, c'est « l'article sur une carte postale ». Le plan doit être respecté. Aucune critique ou appréciation ne doit apparaître, c'est une épreuve de synthèse. Ce résumé est important car c'est la partie la plus lue d'un article, secondairement lorsque sont réalisées des recherches bibliographiques.

### *A) La structuration du résumé :*

La plus probable sera : objectifs, matériels et méthodes, résultats, conclusion.

L'objectif est habituellement écrit à la fin de l'introduction de l'article.

La conclusion est habituellement écrite au début de la discussion de l'article.

Matériels et méthodes doit être court, reprendre les « mots clefs » de la méthodologie.

Les résultats méritent d'être hypertrophiés.

Pour « matériels et méthodes » et pour « résultats », il faut avoir un plan :

- matériels et méthodes :      sur qui ou quoi porte l'étude,  
   sélection, randomisation ?  
   critères diagnostiques  
   évaluation des résultats  
   comment a-t'on procédé ?

- résultats : les résultats essentiels à la discussion.

### *B) Quelques conseils concernant le style :*

- faire des phrases simples,
- courtes,
- avec des mots simples,
- hypertrophier les résultats,
- faire des choix,
- ne pas répéter le titre,
- pas d'abréviation (ou expliquer),
- pas de référence, respecter le temps des verbes.

### *C) Structuration du résumé :*

Problématique ou objectif : 70 mots

Matériels et Méthodes : 50 mots

Résultats : 100 mots

Conclusion : 30 mots.

Ces chiffres sont donnés à titre indicatif et permettent de comprendre comment doit être présenté le résumé.

Vous trouverez ci-joint les recommandations officielles pour l'écriture de ce résumé.

Résumé en 250 mots maximum comportant : Objectif, Matériel et Méthodes, Résultats, Discussion (le tableau comporte 254 cases, les têtes de paragraphe doivent figurer chacune dans une case)

1 à 6						
7 à 12						
13 à 18						
19 à 24						
25 à 30						
31 à 36						
37 à 42						
43 à 48						
49 à 54						
55 à 60						
61 à 66						
67 à 72						
73 à 78						
79 à 84						
85 à 90						
91 à 96						
97 à 102						
103 à 108						
109 à 114						
115 à 120						
121 à 126						
127 à 132						
133 à 138						
139 à 144						
145 à 150						
151 à 156						
157 à 162						
163 à 168						
169 à 174						
175 à 180						
181 à 186						
187 à 192						
193 à 198						
199 à 204						
205 à 210						
211 à 216						
217 à 222						
223 à 228						
229 à 234						
235 à 240						
241 à 246						
247 à 252						
253 à 254						

Comment seront comptés les mots lors de l'épreuve  
de lecture critique d'article (LCA) (2009)

1. **Comptent comme un mot (une case) :**

un mot :

- ✓ simple ou composé avec ou sans tiret (exemple: globulines, gamma globulines,  $\alpha$  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 ( $\alpha$ ,  $\beta$ ...).

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.m<sup>2</sup>, 26 m/s).

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

## CNCI-Conseil Scientifique Médecine

### Liste des abréviations autorisées (dans les questions et les réponses)

Toute autre abréviation peut être utilisée dans un cas clinique si et seulement si l'abréviation est décrite en début de texte.

Liste revue le 16/12/2004 par le Conseil Scientifique Médecine du CNCI

#### A

- ACE : antigène carcino-embryonnaire
- ACTH : adrenocorticotrop hormone (corticotrophine, hormone corticotrope hypophysaire)
- ADH : antidiuretic hormone (hormone antidiurétique, vasopressine)
- ADN : acide désoxyribonucléique
- AINS : anti inflammatoire non stéroïdien
- ALAT : alanine amino transferase (TGP)
- ALD : affection de longue durée
- AMM : autorisation de mise sur le marché
- AMPc : AMP cyclique
- ANCA : antineutrophilic cytoplasmic antibody (anticorps anti cytoplasme des polynucléaires)
- APGAR : american pediatric gross assessment record
- APUD : amine precursor uptake and decarboxylation (groupe de cellules captant et décarboxylant des précurseurs d'amines)
- ARN : acide ribonucléique
- ARNm : ARN messenger
- ASA : classification du risque opératoire de l'american society of anesthesiologist
- ASAT : aspartate amino transférase (TGO)
- ASLO : antistreptolysine O
- ATP : adénosine triphosphate
- AVC : accident vasculaire cérébral
- aVf, aVL, aVr : dérivations électrographiques unipolaires
- AVK : anti-vitamine K

#### B

- BCG : bacille de Calmette et Guérin
- BK : bacille de Koch
- BPCO : broncho pneumopathie chronique obstructive

#### C

- CCMH : concentration corpusculaire moyenne en hémoglobine
- CEC : circulation extra corporelle
- CGMH : concentration globulaire moyenne en hémoglobine
- CIVD : coagulation intra-vasculaire disséminée
- CK : créatine kinase
- CMV : cytomégalovirus
- CO : monoxyde de carbone
- CO<sub>2</sub> : dioxyde de carbone
- CPK : créatine phosphokinase
- CPK-BB : créatine phosphokinase iso enzyme BB
- CPK-MB : créatine phosphokinase iso enzyme MB
- CPK-MM : créatine phosphokinase iso enzyme MM
- CRH : corticotropin releasing hormone (hormone de libération de l'hormone corticotrope)
- CRP : C reactive protein (protéine C réactive)

#### D

- DCI : dénomination commune internationale
- DHEA : déhydroépiandrostérone
- DOPA : dihydroxyphénylalanine

**E**

- EBNA : Epstein-Barr nuclear antigen
- EBV : Epstein-Barr virus
- ECBU : examen cytobactériologique des urines
- ECG : électrocardiogramme
- ECHO virus : enteric cytopathogenic human orphan virus
- EEG : électroencéphalogramme
- EFR : épreuve fonctionnelle respiratoire
- ELISA : enzyme-linked immunosorbent assay
- EMG : électromyographie

**F**

- FiO<sub>2</sub> : fraction inspirée d'oxygène
- FSH : follicle stimulating hormone (hormone folliculostimulante)

**G**

- gamma-GT : gamma-glutamyl transférase
- GB : globule blanc
- GH : growth hormone (hormone somatotrope)
- GH-RH : GH-releasing hormone (hormone activatrice de l'hormone de croissance)
- GR : globule rouge
- GVH : graft versus host (réaction du greffon contre l'hôte)

**H**

- Hb : hémoglobine
- HbA1C : hémoglobine glyquée
- HbO<sub>2</sub> : oxyhémoglobine
- HBPM : héparine de bas poids moléculaire
- HCG : human chorionic gonadotrophin (gonadotrophine chorionique)
- HDL : high density lipoproteins (lipoprotéines de haute densité)
- HELLP : hemolysis + elevated liver enzymes + low platelets
- HLA : human leucocyte antigen (antigène d'histocompatibilité)
- HPV : human papillomavirus
- HTLV : human T cell leukemia/lymphoma virus (virus humain T lymphotrope)

**I**

- IDR : intradermoréaction
- IEC : inhibiteur de l'enzyme de conversion
- Ig : immunoglobulines
- IGF : insulin-like growth factor (somatomédine)
- IMAO : inhibiteur de la mono amine oxydase
- INR : international normalized ratio
- IRM : imagerie par résonance magnétique
- (exit: ITT : incapacité temporaire totale)(ambiguïté=>supprimée le 17/3/2005)
- IV : intraveineuse

**L**

- LCR : liquide céphalorachidien
- LDH : lactate déshydrogénase
- LDL : low density lipoprotein (lipoprotéine de faible densité)
- LH : luteinizing hormone (hormone lutéinisante)
- LHRH : luteinizing hormone releasing hormone (gonadolibérine)

**M**

- MALT : mucous associated lymphoid tissue (tissu lymphoïde associé aux muqueuses)
- MNI : mononucléose infectieuse
- MST : maladie sexuellement transmissible

**N**

- NFS : numération formule sanguine
- NK : natural killer (lymphocyte)

**O**

- OAP : œdème aigu du poumon
- OMS : organisation mondiale de la santé
- ORL : oto-rhino-laryngologie

**P**

- PAN : périartérite noueuse
- PaO<sub>2</sub> : pression artérielle partielle en oxygène
- PaCO<sub>2</sub> : pression artérielle partielle en dioxyde de carbone
- PCR : polymérase chain-reaction
- PDF : produits de dégradation de la fibrine
- PDGF : platelet-derived growth factor (facteur de croissance des plaquettes)
- PMI : protection maternelle et infantile
- PNB : polynucléaires basophiles
- PNE : polynucléaires éosinophiles
- PNN : polynucléaires neutrophiles
- PSA : prostatic specific antigen

**Q**

- QI : quotient intellectuel
- QRS : complexe QRS
- QSP : quantité suffisante pour
- QT : segment QT

**R**

- RAST : radio allerge sorbent test (dosage radio immunologique des IgE spécifiques d'un allergène)
- Rh : rhésus
- RMN : résonance magnétique nucléaire

**S**

- SIDA : syndrome d'immunodéficience acquise
- SRAS : syndrome respiratoire aigu sévère

**T**

- T3 : triiodothyronine
- T4 : thyroxine, tétraiodothyronine
- TCA : temps de céphaline activée
- TCK : temps de céphaline kaolin
- TCMH : teneur corpusculaire moyenne en hémoglobine
- TDM : tomodensitométrie
- TEP : tomographie par émission de positon
- TGMH : teneur globulaire moyenne en hémoglobine
- TGO : transaminase glutamo oxaloacétique
- TGP : transaminase glutamo pyruvique
- TNM : classification tumor nodes metastasis (tumeur primitive, adénopathies régionales, métastases)
- TP : taux de prothrombine
- TPHA : treponema pallidum haemagglutination assay
- TRH : thyrotropin releasing hormone (protiréline)
- TSH : thyroid stimulating hormone (thyroestimuline)

**U**

- UI : unité internationale
- UIV : urographie intraveineuse

**V**

- VDRL : veneral disease research laboratory (réaction d'agglutination syphilitique)
- VEMS : volume expiratoire maximum par seconde
- VGM : volume globulaire moyen
- VIH : virus de l'immunodéficience humaine (HIV, virus du SIDA)
- VLDL : very low density lipoproteins (lipoprotéines de très faible densité)
- VS : vitesse de sédimentation

Il est important de surligner, lors de la lecture de l'article, et probablement de couleur différente, les phrases qui peuvent être importantes pour chaque section du résumé. Il faut garder le temps nécessaire pour pouvoir écrire correctement ce résumé.

Il paraît également intéressant, puisqu'il faut rapporter dans la grille chacun des titres de section (objectifs, méthodes, résultats, discussion), de les surligner ou de les souligner pour que le correcteur voit bien la différence entre chaque section.

Enfin, la principale difficulté réside sur la définition du mot. En cas d'ambiguïté, c'est le jury qui reste souverain dans cette définition.



**CHAPITRE V :**  
**LECTURE CRITIQUE D'UNE ANALYSE STATISTIQUE**  
**METHODES ET INTERPRETATION**

## Plan

Introduction

I- Rappel sur la lecture critique de la méthodologie d'une étude

II- Plan habituel de la section méthodologie statistique

III- Eléments principaux de la lecture critique statistique

Conclusion

Tableau des tests

## INTRODUCTION :

Dans les objectifs pédagogiques de l'épreuve de lecture critique d'un article médical scientifique mis en place au concours de l'internat, plusieurs objectifs se rapportent spécifiquement à la méthodologie statistique :

- analyser la technique de randomisation,
- vérifier que les analyses statistiques sont cohérentes avec le projet de travail et discuter leur interprétation,
- vérifier la présentation des indices de dispersion des données numériques et celles de l'imprécision des estimations,
- discuter la signification statistique des résultats.

Le présent chapitre se propose d'aider l'étudiant à réaliser ces objectifs et rappelle si cela était nécessaire l'ensemble des objectifs pédagogiques établis par le centre national des concours d'internat (CNCI).

La lecture « critique » n'implique pas forcément une critique péjorative, elle se conçoit plutôt comme l'appréciation (en termes positifs ou négatifs) de la qualité scientifique d'une étude.

Après avoir fait quelques rappels sur la lecture critique de la méthodologie d'une étude qui a déjà été abordée dans le chapitre précédent, en insistant sur la méthodologie des essais thérapeutiques, nous aborderons successivement le plan habituel de la section statistique des méthodes d'un article original, puis quelles sont les questions principales à se poser pour critiquer les méthodes statistiques utilisées et l'interprétation qui en est faite. Nous essaierons de dégager les pièges les plus fréquents que l'on peut rencontrer dans cet exercice souvent délicat.

## **I- RAPPELS SUR LA LECTURE CRITIQUE DE LA METHODOLOGIE D'UNE ETUDE :**

Des critères généraux de qualité de la méthodologie d'une étude originale peuvent être posés :

- l'objectif principal d'une étude doit être clair et une seule hypothèse doit être posée a priori,
- le critère principal de jugement découle de l'objectif et doit être unique
- le schéma d'étude doit être adapté à la question posée,
- les procédures de sélection des patients doivent être correctes,
- l'étude doit être conduite correctement.

Selon les schémas d'étude, ces critères peuvent être plus ou moins importants. De plus, il existe des critères spécifiques de qualité de la méthodologie en fonction du schéma d'étude. Ces

critères de qualité peuvent être retrouvés dans les grilles de lecture critique qui sont un outil précieux pour l'analyse d'articles (voir chapitre précédent).

Il convient de déterminer si l'article apporte des éléments originaux par rapport à la littérature existante. En particulier, l'étude peut être plus longue, ayant inclus plus de sujets, ou ayant pris en compte plus de paramètres que les études précédentes sur le même thème. La méthodologie de cette étude peut être plus rigoureuse, en particulier si elle a pris en compte les critiques méthodologiques spécifiques aux études antérieures. L'étude peut être également originale si elle a inclus une population différente des études antérieures, par exemple différents groupes d'âge, ou des pathologies différentes des études antérieures. Enfin, l'étude peut également être innovante en termes d'importance clinique, pour le patient, pour la société, et notamment pour les décideurs.

Les différents objectifs pédagogiques listés par le CNCI et concernant ce chapitre sont les suivants (rappel du chapitre I) :

1) Identifier :

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

1.2) La « question » étudiée

1) Critiquer la méthodologie : population étudiée :

2.1) Identifier les caractéristiques de la population étudiée

2.2) Analyser les modalités de sélection des sujets, critères d'inclusion et critères d'exclusion

2.3) Analyser la technique de randomisation, le cas échéant.

2.4) Discuter la comparabilité des groupes soumis à la comparaison

2.5) Discuter le choix des effectifs étudiés et leur cohérence dans la totalité de l'article

Comment les sujets ont-ils été recrutés ?

Un **recrutement exhaustif** sur une période de temps donné, ou dans un lieu donné est correct sauf phénomène saisonnier pouvant être lié à la question de recherche. Un **tirage au sort** parmi l'ensemble des sujets éligibles est bien sûr l'idéal.

Qui a été inclus dans cette étude ? la plupart des études excluent les patients avec des comorbidités, les patients qui peuvent poser un problème de suivi ou de compréhension de l'essai, du langage, à la probabilité de perte de vue du sujet ; d'autres excluent les sujets avec des médicaments co-existants, etc ... Même si cette approche est valide scientifiquement, il est clair que plus ces critères d'exclusion sont importants, plus on s'éloigne de la pratique clinique quotidienne et plus les résultats de cette étude seront difficiles à appliquer dans d'autres populations que la population de l'essai. Un essai randomisé peut par exemple être restreint aux patients ayant certaines formes de maladie, les plus sévères par exemple, ce qui aura bien sûr d'importantes implications pratiques si l'on veut appliquer les résultats de l'essai à une population différente en pratique médicale quotidienne car les formes cliniques seront certainement plus variées (généralisation des résultats).

Les sujets ont-ils été étudiés dans des situations d'observation habituelles ? ici encore, les écarts à la pratique quotidienne n'invalideront pas les résultats scientifiques de l'étude, mais risquent de rendre difficile une extrapolation des résultats de l'essai à d'autres populations. Par exemple, certains essais obligent les patients à être hospitalisés, ce qui peut ne pas être le cas habituel dans la pratique courante ; les notices d'information des essais thérapeutiques apportent une information beaucoup plus étendue et complexe que celle qui est donnée en pratique quotidienne au patient ; certains équipements peuvent être utilisés alors qu'ils ne le seront pas facilement en pratique routinière.

3) Critiquer la méthodologie : méthode :

3.1) S'assurer que la méthode employée est cohérente avec le projet du travail et quelle est effectivement susceptible d'apporter « une » réponse à la question posée dans l'introduction.

*Le protocole est-il clairement et suffisamment expliqué dans l'article et les mesures effectuées valides ?*

L'article doit en particulier préciser quelle intervention spécifique a été réalisée et à quoi elle a été comparée. Ceci paraît évident, mais souvent les paragraphes méthodes des articles originaux ne sont pas suffisamment précis. Les informations données ne sont souvent pas suffisantes à propos des sujets ou de la prise en charge réalisée.

Les « nombres d'or » de l'étude ont-ils été précisés dans les méthodes ?

Les méthodes d'un essai thérapeutique doivent comporter trois nombres essentiels :

- **le calcul du nombre de sujets nécessaires,**
- **la durée du suivi,**
- **l'exhaustivité du suivi.**

Nous reviendrons plus loin sur le paragraphe du nombre de sujets nécessaires.

La durée de suivi est importante, même si le nombre de sujets nécessaires est adéquat. Des données précises doivent être fournies par les auteurs, en particulier sur la médiane de suivi.

L'exhaustivité du suivi et notamment le nombre de sujets qui ont été perdus de vue et pourquoi, doit être précisée. Le diagramme de flux des patients est un schéma important recommandé par les règles de rédaction d'un article rapportant un essai thérapeutique, règles qui ont été mises au point par plusieurs journaux internationaux (CONSORT Statement).

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

3.3) Vérifier les règles d'éthique

4) Critiquer la présentation des résultats :

4.1) Analyser la présentation, la précision et la lisibilité des tableaux et des figures, leur cohérence avec le texte et leur utilité.

4.2) Vérifier la présentation des indices de dispersion (valeurs extrêmes, quantiles, écarts types) des données numériques, et celle de la précision des estimations (intervalle de confiance, variance, ...)

5) Critiquer l'analyse des résultats et de la discussion :

5.1) Discuter la nature et la précision des critères de jugement réalisés

Les auteurs doivent préciser quel critère de jugement a été utilisé. Un seul critère de jugement **principal** doit être indiqué et c'est sur ce critère de jugement que les auteurs doivent baser leurs résultats et leur interprétation. Encore faut-il que ce critère de jugement soit pertinent, par exemple en terme de survie, de qualité de vie ou de meilleure tolérance. Des marqueurs indirects de l'efficacité d'un produit (en anglais surrogate markers) peuvent être utilisés (exemples : marqueurs biologiques...) mais ne devraient pas constituer la base de tout le raisonnement des auteurs, surtout si on souhaite utiliser les résultats de l'article dans une large pratique. Les critères de jugement doivent être objectivement valides, c'est-à-dire sensibles et spécifiques. Ceci est en particulier très important lorsque les mesures ont concerné des symptômes, des effets psychologiques ou sociaux.

5.2) 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

*Tout a-t-il été fait dans l'étude pour éviter les biais ou les minimiser ?*

On distingue essentiellement trois grands types de biais :

- les biais de sélection
- les biais d'information
- les biais de confusion

**Les biais de sélection** sont liés au mode de sélection des sujets (échantillonnage) ou à leur exclusion secondaire (perdus de vue). Ces biais posent le problème de la non représentativité de l'échantillon. Les auteurs ont-ils sélectionné un échantillon aléatoire ? Le nombre de non répondants est-il suffisamment faible et ceux-ci ressemblent-ils aux répondants (biais d'exclusion) ? Les sujets recrutés sont-ils particuliers : patients hospitalisés, sujets auto-sélectionnés en fonction de l'objectif de l'étude (survie sélective ou cas uniquement prévalents par exemple) ... Y a-t-il eu beaucoup de sujets perdus de vue ou exclus secondairement (attrition) qui pourrait mettre en cause la comparabilité initiale ?

**Les biais d'information** (on trouve parfois biais de détection, biais de mesure, ou biais de classement en fonction des contextes) sont dus à l'utilisation de fausses informations. Ces biais sont essentiellement liés au mode de recueil des informations sur le facteur d'exposition ou sur les diagnostics et aux modalités d'interrogatoire des sujets : non connaissance de leur état de santé, auto-questionnaire déclaratif, mensonge (biais de prévarication), refus de réponse, non compréhension des questions, problème de mémoire (biais de rappel) ... Ils peuvent être également liés à des problèmes de codage des informations avant leur exploitation statistique. La motivation de l'enquêteur ou sa connaissance de l'objectif de l'enquête ou du statut de la personne enquêtée (cas ou témoin par exemple, ou exposé non exposé) peut influencer sur les réponses des sujets ou l'interprétation éventuelle de la réponse par l'enquêteur. Un biais d'information peut être non différentiel (l'erreur est la même dans les différents groupes de l'étude : par exemple peser tous les sujets avec une balance imprécise) ou différentiel (l'erreur est plus importante ou n'existe que dans un des groupes : par exemple peser seulement les cas avec une balance imprécise et les témoins avec une balance très précise). Evidemment, une erreur différentielle est plus grave qu'une erreur non différentielle.

**Les biais de confusion** sont dus à la non prise en compte d'un ou plusieurs facteurs de confusion. Un facteur de confusion est un facteur lié statistiquement à la fois au facteur de risque et

à la maladie. Exemple : si on cherche un lien éventuel entre le sexe et le cancer bronchique, il faut prendre en compte le tabagisme (lié au sexe car plus fréquent chez les hommes) et lié évidemment au cancer bronchique... On peut neutraliser l'effet d'un facteur de confusion en utilisant différentes techniques : avant le début de l'étude, on peut appairer les individus (on les rend semblables sur certains éléments : par exemple, appariement des cas et des témoins sur l'âge et le sexe : pour tout malade d'âge et de sexe donné, on lui associe un témoin de même âge ou à +/- 5 ans et de même sexe). On peut également ajuster après le début de l'étude en utilisant des méthodes statistiques d'analyse multivariée. Le plus souvent, le mot « confusion » n'apparaît pas comme tel dans les articles ; les facteurs de confusion sont les variables sur lesquelles on apparie ou on ajuste à l'analyse multivariée.

**En lecture critique**, il faut s'assurer que les auteurs discutent bien des biais de leur étude et de l'influence potentielle de ces biais sur leurs conclusions. Aucune étude n'est parfaite. Il y a donc toujours de quoi discuter. Il faut également pouvoir décrire la façon dont les auteurs ont pris en compte d'éventuels facteurs de confusion (ajustement par analyse multivariée la plupart du temps).

Un biais résulte de tout élément qui va influencer les conclusions de l'étude, distordre les comparaisons entre les groupes et entraîner des conclusions erronées. Quelque soit le protocole de l'étude (essai randomisé, essai non randomisé, étude de cohorte ou étude cas-témoins), tout doit être fait pour que **les groupes comparés soient le plus possible semblables**, excepté pour la différence recherchée. Par exemple, chacun des groupes doit recevoir les mêmes explications sur l'étude, avoir les mêmes contacts avec les professionnels de santé, être examiné le même nombre de fois en utilisant les mêmes critères de jugement.

Dans un essai thérapeutique randomisé, le biais est en théorie évité en sélectionnant un échantillon de participants dans une population spécifique, puis en allouant par tirage au sort les différents traitements étudiés à différents groupes. Pour ce qui est des essais non randomisés, la comparaison des différences à l'inclusion dans l'étude doit être particulièrement fine, afin de permettre aux lecteurs de juger si ces différences sont suffisamment importantes pour invalider des différences pouvant être rapportées à l'effet de l'intervention. L'absence de randomisation doit rendre extrêmement prudente l'interprétation des résultats d'une étude. Dans une étude épidémiologique d'observation, que ce soit cohorte ou cas-témoins, la sélection de groupes comparables est le point essentiel. Les différences peuvent parfois être contrôlées au moment de l'analyse statistique, grâce à des ajustements complexes, mais cela doit rester le dernier recours.

Les procédures doivent être les mêmes dans les deux groupes. Une différence pourrait conduire à un biais d'exécution (qui entre dans le cadre d'un biais de confusion). Les évaluations



doivent être faites le plus possible **en insu (aveugle)**. Le double insu doit être une règle dans les essais randomisés et s'il n'a pas été utilisé, l'auteur doit fournir des explications détaillées sur les raisons de cette absence et sur ses implications pour les résultats de l'étude. S'il n'a pas été utilisé, le critère de jugement doit être le plus objectif possible (examen biologique, qui lui peut être éventuellement réalisé ou interprété en insu...).

Enfin, l'analyse de l'essai doit être effectuée en intention de traiter. Ceci veut dire que les patients doivent rester dans leurs groupes de randomisation initial au moment de l'analyse, même si des difficultés sont survenues ultérieurement, c'est-à-dire s'ils n'ont pas pris leur traitement correctement, si des réactions indésirables sont survenues obligeant à l'arrêt du traitement, si d'autres raisons cliniques ont obligé à arrêter l'essai chez un patient, etc ...

**Les seuls résultats valides d'un essai thérapeutique randomisé sont constitués par les résultats de l'analyse en intention de traiter.**

- 5.3) Vérifier la logique de la discussion et sa structure. Reconnaître ce qui relève des données de la littérature et ce qui est opinion personnelle de l'auteur
- 5.4) Discuter la signification statistique des résultats
- 5.5) Discuter la signification clinique des résultats
- 5.6) Vérifier que les résultats offrent une réponse à la question annoncée
- 5.7) Vérifier que les conclusions sont justifiées par les résultats
- 5.8) Indiquer le niveau de preuve de l'étude (grille de l'HAS)

## **II- PLAN HABITUEL DE LA SECTION METHODOLOGIE STATISTIQUE :**

Dans la partie méthodes d'un article original, doit être individualisée une partie méthodologie statistique. Cette partie ne se conçoit qu'à la fin d'une méthodologie complète, comportant des détails sur la variable dépendante (critère de jugement principal), la variable indépendante principale et les autres paramètres mesurés. Elle ne se conçoit également qu'après des détails sur le recueil des données, sur leur circuit (contrôle, saisie). Le paragraphe « nombre de sujets nécessaires » peut être intégré dans cette partie analyse statistique, mais peut être également séparé. L'analyse statistique doit être planifiée a priori, ce qui est évidemment difficile à juger au stade de la rédaction du papier. Toutefois, l'apparition d'analyses en sous-groupes dans les résultats doit rendre le lecteur prudent.

La partie méthodologie statistique doit pouvoir répondre à deux questions :

- quels calculs et quels tests ont été réalisés ?
- pour aboutir à quels types de résultats ?

Cette partie doit être en adéquation avec le schéma d'étude, le type de variables mesurées, le rôle des variables principales (dépendantes ou indépendantes) et leur niveau de mesure.

Le plan général habituel d'une partie analyse statistique est le suivant :

1. Modalités d'analyse descriptive de l'échantillon : les auteurs doivent préciser comment sont exprimés les résultats descriptifs des variables qualitatives (pourcentages habituellement), ou des variables quantitatives (habituellement moyennes  $\pm$  écart-types). Pour les variables quantitatives, d'autres paramètres peuvent être également donnés comme les médianes (valeur qui sépare la distribution en deux parties égales), les extrêmes (souvent important pour se rendre compte de l'homogénéité de la population). L'écart-type est encore appelé déviation standard ou *standard déviation* en anglais. Parfois, d'autres mesures de dispersion sont utilisées comme l'erreur standard ou *standard error* en anglais, qui est en fait l'écart type de la moyenne et qui est calculé par l'écart type divisé par la racine carrée de l'effectif. Il est alors facile de comprendre que cette erreur standard est toujours plus faible que l'écart-type, ce qui peut donner aux lecteurs une fausse impression de grande précision des résultats de l'étude, alors qu'il s'agit simplement d'une astuce de présentation. Ceci peut être vrai pour les chiffres donnés dans l'article, mais également pour les représentations graphiques. Il est donc de bonne pratique que les auteurs précisent s'ils ont exprimé leurs résultats sous la forme de moyenne  $\pm$  écart type ou sous la forme de moyenne  $\pm$  erreur standard.
2. Modalités de traitement des non réponses, ou des valeurs manquantes en général : dans certaines études, les valeurs manquantes peuvent être nombreuses pour différentes raisons : perdus de vue, absence de réponse, erreur de saisie, ... Les auteurs devraient préciser la fréquence de leur survenue. Cette fréquence pouvant varier pour chaque paramètre mesuré, elle n'est souvent pas détaillée mais, au minimum, les auteurs doivent donner une idée de l'ampleur de ces valeurs manquantes et de ce qui a été fait pour les prendre en compte ou non dans les analyses. Dans certaines études (cohortes par exemple), des traitements spécifiques de ces valeurs manquantes peuvent être mis en œuvre.

3. Analyse du critère principal de jugement : le test le plus important est bien sûr celui du critère principal de jugement entre les groupes comparés. Le test statistique doit être précisé et justifié. Un seul type de test doit être réalisé à ce niveau. Ce test doit être adapté à la question posée (voir plus loin). Dans cette partie de méthodologie statistique, seul le test peut être précisé, mais il est bien évident que dans la partie résultats, tout test statistique doit s'accompagner des valeurs qui ont été comparées. Dans la partie résultat, le test doit être résumé par le degré de significativité « p ». Un « p » sans les valeurs qui ont été comparées n'a aucune signification et ne peut être réellement interprété.
4. Analyse des critères secondaires : les autres critères de jugement (idéalement peu nombreux) doivent être également comparés et les tests statistiques adaptés doivent être précisés.
5. Prise en compte d'éventuelles co-variables : l'influence de co-variables sur la question principale de jugement peut être recherchée, soit co-variable par co-variable (analyse univariée), soit de façon multivariée. Le type d'analyse multivariée doit être alors précisé, ainsi que la procédure réalisée : sélection des variables incluses dans le modèle initial (par exemple utilisation du seuil de significativité 0,25 ou 0,20 en analyse univariée pour ne pas négliger des variables importantes) puis sélection progressive des variables pour aboutir au modèle final (le seuil de signification est alors de 0,05, comme pour toute autre analyse).
6. Analyses inter-co-variables : celles-ci doivent être limitées car il ne s'agissait pas de la question de recherche. Elles peuvent simplement être présentées comme génératrices d'hypothèses nouvelles pour d'éventuelles études ultérieures.
7. Autres analyses : les auteurs peuvent présenter, à la fin de cette partie de méthodologie statistique, d'autres analyses réalisées comme des analyses de concordance entre plusieurs observateurs ou entre plusieurs techniques de mesure, les méthodes d'estimation de la précision de certaines mesures (intervalles de confiance), etc ...

Ce plan est évidemment un plan classique mais peut varier selon les schémas d'étude et selon les habitudes des auteurs. Ce paragraphe de méthodologie statistique doit être rédigé dans des termes précis, en style scientifique, de façon suffisamment détaillée. Ce paragraphe doit comporter

le seuil de significativité qui a été choisi, soit globalement pour l'ensemble des analyses statistiques (habituellement **seuil à 0,05**, les auteurs se laissent 5% de risque d'erreur de conclure à une différence statistiquement significative alors qu'elle n'existe pas réellement). L'inclusion des variables dans le modèle initial d'une analyse multivariée peut être faite en dessous d'un seuil de 0,25 mais le modèle final tiendra compte du seuil à 0,05.

Ce paragraphe doit également préciser si une seule analyse statistique du critère principal de jugement a été réalisée pour l'ensemble de l'étude ou si des analyses intermédiaires ont été prévues dès le début du protocole. Auquel cas, un ajustement du risque doit avoir été prévu dès le calcul du nombre de sujets nécessaires, afin de ne pas dépasser un risque d'erreur global de l'analyse statistique sur l'ensemble de l'étude. Ce même problème existe également pour les comparaisons multiples qui doivent avoir incité les auteurs à ajuster ces risques d'erreurs.

Enfin, et cela n'est pas souvent précisé dans les paragraphes de méthodologie statistique, les conditions d'application des tests doivent avoir été vérifiées. Il est bien évident qu'il est nécessaire de faire confiance aux auteurs pour ce qui est des tests habituels, mais lorsque les analyses sont plus complexes, les auteurs doivent détailler dans leur article les méthodes qu'ils ont mis en œuvre pour vérifier les hypothèses sous jacentes aux tests réalisés.

### **III- ELEMENTS PRINCIPAUX DE LA LECTURE CRITIQUE STATISTIQUE :**

#### 1) Randomisation :

La seule manière de rendre comparable deux groupes de sujets avant d'appliquer une intervention spécifique à l'un des groupes est de randomiser, c'est-à-dire de tirer au sort le groupe de traitement dans lequel les sujets se situeront. Toute autre méthode risque de mener à des biais lors des comparaisons statistiques. Les techniques de randomisation sont multiples. Il existe des randomisations simples, des randomisations stratifiées sur certaines variables, des randomisations stratifiées par centre dans les essais multicentriques. La randomisation fait appel à des tables de nombres aléatoires. Une autre technique, qui commence à être utilisée, est la technique de minimisation qui n'est pas une réelle randomisation, mais qui cherche à minimiser les différences entre les variables pronostiques principales entre les deux groupes.

La randomisation la plus utilisée dans un essai thérapeutique est la randomisation équilibrée par blocs (quelques sujets) de taille variable. Par exemple : bloc de taille 4 : AABB (les 2 premiers sujets recevront le traitement A, les deux derniers le traitement B), puis bloc de taille 6 (ABABAB) etc... Elle assure que les différents groupes auront un effectif identique à des moments précis (à la fin des blocs) au fur et à mesure des inclusions. De plus le caractère variable de la taille des blocs

permet d'éviter toute déduction du bras (groupe de traitement) dans lequel est randomisé un patient sur la base de la connaissance du traitement alloué aux précédents sujets.

Même en cas de randomisation satisfaisante, les auteurs doivent donner une table montrant les caractéristiques de base des groupes étudiés. Cette table doit montrer que les groupes sont similaires pour les variables pronostiques importantes. Théoriquement, si la randomisation a été correctement réalisée, aucune différence ne sera mise en évidence. Toutefois, si d'importantes différences sont retrouvées à ce stade, même dues au hasard, cela peut poser un problème d'interprétation dans les résultats. Il faut donc alors voir si les auteurs ont pris en compte ces différences et ont par exemple réalisé des ajustements dans les analyses statistiques ultérieures.

## 2) Nombre de sujets nécessaires :

Le calcul de la taille de l'échantillon prend en compte plusieurs éléments. Ce calcul doit se faire a priori, au moment de la planification de l'étude et doit être adapté à la question posée. Habituellement, il est placé dans la section méthodes, avant la description de l'analyse statistique. L'effet considéré comme cliniquement significatif peut être indiqué en premier, suivi des niveaux de risque statistique que les auteurs se sont fixés (risque alpha (risque consenti de mettre en évidence une différence qui n'existe pas, habituellement 5%) et risque beta (risque consenti de ne pas mettre en évidence une différence qui existe, habituellement 20%), (ou puissance c'est à dire  $1 - \beta$ ). Le caractère unilatéral ou bilatéral du test influe également sur ce nombre de sujets (voir point 4 ci-dessous). Enfin, des données concernant la variabilité des phénomènes mesurés peuvent être indiquées.

Exemple : le nombre de sujets nécessaire a été calculé pour mettre en évidence une réduction du critère de jugement principal de « x », en se fixant des risques alpha et bêta de y et z, pour un test bilatéral (ou unilatéral) et un écart-type du critère de jugement principal « s ». Ce nombre de sujets est de « n ».

Parfois, le calcul de la taille des échantillons n'a pas pu être fait a priori, à cause d'un manque de données précises sur la variabilité du phénomène étudié. Il arrive parfois que l'on réalise une étude préliminaire pour obtenir ces données, ou que l'on réajuste le nombre de sujets prévus initialement en fonction des données obtenues sur les premiers sujets inclus. Cela doit apparaître clairement dans la section méthodes et la discussion de l'article doit évoquer cette limite et ses conséquences éventuelles.

### 3) Les tests statistiques étaient-ils appropriés ?

Le choix d'un test statistique reste du domaine du statisticien ! Cependant l'étudiant doit connaître le nom de tests statistiques usuels et à quoi ils servent. L'important est que ce test soit adapté au type de données disponibles. Les tests statistiques sont différents si l'on travaille sur des données qualitatives, sur une comparaison entre données qualitatives et quantitatives, ou uniquement sur des données quantitatives (tableau I). Une variable quantitative est décrite habituellement par une moyenne couplée éventuellement à un écart type. Une variable qualitative s'exprime en pourcentage.

**Tableau I : Niveau de mesure des variables et choix des méthodes statistiques**  
(cas des comparaisons non appariées)

<i>Mesure des variables</i>		<i>Méthodes statistiques possibles</i>	
<i>Dépendante</i>	<i>Indépendante</i>	<i>Paramètre estimé</i>	<i>Test statistique</i>
<b>Dichotomique</b>	Dichotomique	Proportions ; risques relatifs	$\chi^2$ ; test exact de Fisher
	Multinomiale	Proportions ; risques relatifs	$\chi^2$
	Ordinale	Proportions ; risques relatifs	$\chi^2$ ; test $\chi^2$ de tendance
	Quantitative	Moyennes ; médianes	Test $t$ ; tests de Mann-Whitney
	Variées	Régression logistique ; analyse descriptive multivariée ;	
<b>Multinomiale</b>	Variées	Régression logistique ; analyse descriptive multivariée ;	
<b>Ordinale</b>	Multiples	Régression logistique ; analyse descriptive multivariée ;	
<b>Quantitative</b>	Dichotomique	Moyennes ; médianes	test $t$ ; tests de Mann-Whitney
	Multinomiale	Moyennes ; médianes	Analyse de variance
	Ordinale	Moyennes ; médianes	Analyse de variance
	Quantitative	Coefficients de corrélation ou régression	Régression linéaire simple
	Variées	Régression linéaire multiple ; modèles linéaires généralisés	
<b>Temps de survenue d'un événement</b>	Variées	Modèles de survie (Kaplan-Meier, Cox, ...)	

Liste non exhaustive ; le choix des méthodes dépend aussi du type de question posée et des conditions d'application des tests. Notes : Variable dichotomique = qualitative binaire ; multinomiale = qualitative nominale à plus de deux classes ; ordinale = qualitative à plus de deux classes ordonnées.

Tiré de : Lecture Critique et Rédaction Médicale Scientifique : comment lire, rédiger et publier une étude clinique ou épidémiologique. LR Salmi, éd. Elsevier

Très schématiquement :

- S'il s'agit de comparer des pourcentages entre deux ou plusieurs groupes de sujets (deux variables qualitatives), le test approprié est un  $\chi^2$ .
- S'il s'agit de comparer une variable quantitative entre deux groupes, c'est-à-dire comparer deux moyennes (les groupes étant définis par les niveaux d'une variable qualitative binaire), le test approprié est un test « t » de Student (test paramétrique) ou un test de Mann-Whitney (test non paramétrique).
- Dans le cas de comparaison de plus de deux moyennes, les tests correspondants sont une analyse de variance (test paramétrique) ou un test de Kruskal-Wallis (test non paramétrique).
- S'il s'agit d'analyser si une variable quantitative évolue en lien avec une autre variable quantitative, il s'agit de corrélation. Les tests appropriés sont alors le calcul du coefficient de corrélation linéaire « r » puis son test (effectué par transformation en t de Student) (test paramétrique) ou le test des rangs de Spearman (test non paramétrique).

Ces tests sont les plus fréquents. D'autres situations peuvent toutefois être rencontrées, comme par exemple la situation des tests appariés : il s'agit de tests réalisés chez les mêmes sujets, comme comparer deux moyennes à deux moments différents dans un même groupe de sujets, avant et après un traitement par exemple. Ces tests sont habituellement plus puissants car la variation inter-individuelle est plus faible dans un même groupe de sujets et il est alors plus facile de montrer une différence. Il existe des tests appariés pour chacune des situations que nous venons de voir.

### **Comment choisir entre test paramétrique et test non paramétrique ?**

Les tests paramétriques font l'hypothèse que les données ont une forme particulière de distribution, le plus souvent une distribution normale. Les tests non paramétriques ne font pas ce type d'hypothèse sur la distribution des données. Ils utilisent les rangs des données et non pas les valeurs réelles. En général, les tests paramétriques sont plus puissants que les tests non paramétriques et doivent donc être utilisés si cela est possible.

Décider si les données sont distribuées de façon normale est difficile mais est déterminant pour le choix du test à réaliser.

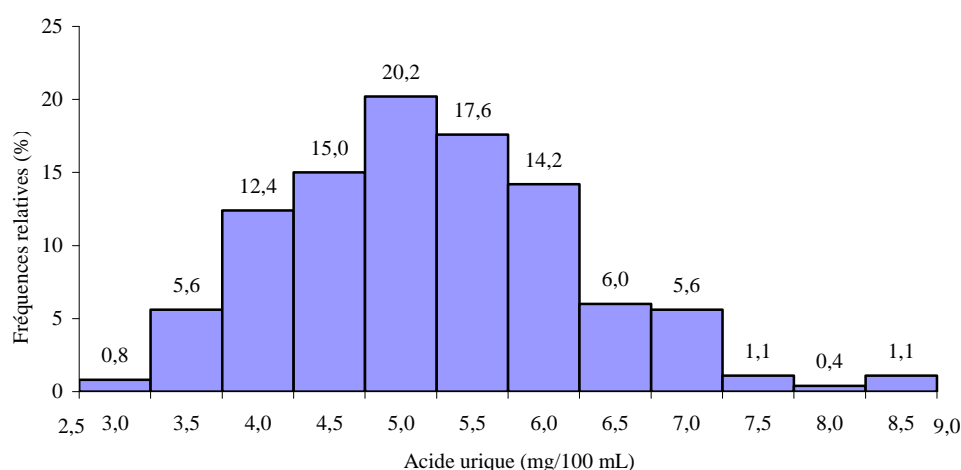
Le simple histogramme de fréquences cumulées d'une variable peut faussement rassurer s'il montre une distribution normale (figure 1). En effet, la distribution de la variable doit être normale dans la population et non pas seulement dans l'échantillon, ce qui est bien entendu difficilement vérifiable. Heureusement, la plupart des données biologiques que l'on utilise en médecine sont distribuées de façon normale autour d'une moyenne dans la population. La plupart des variables peuvent donc être traitées par des tests paramétriques. Toutefois, de nouveaux paramètres à l'étude



dont on ne connaît pas la distribution dans la population ne peuvent être traités par des tests paramétriques que si l'effectif des groupes est suffisamment important (au minimum au-delà de 30 dans chaque groupe). Ce seuil de 30 a été choisi car c'est à partir de ce seuil que l'on peut faire l'approximation que la distribution suit une loi normale, et que dès lors les résultats des tests paramétriques ne seront pas biaisés. Certaines variables dont on sait à l'évidence qu'elles ne suivent pas une loi normale, comme par exemple la durée de séjour dans un service hospitalier, ne doivent pas être traitées par des tests paramétriques, même si le nombre de sujets est important.

**Figure 1 : Histogramme représentant la distribution de la concentration de l'acide urique sérique dans une population de 267 hommes**

*d'après Morton RF, Hebel JR*



Tiré de : Biostatistiques au quotidien. Michel Huguier, Antoine Flahault, éd. Elsevier

Si des auteurs ont utilisé des tests statistiques non usuels, voire obscurs, il faut que l'article justifie leur intérêt et qu'ils soient référencés dans l'article. Les comparaisons statistiques habituelles peuvent être réalisées en utilisant au maximum une douzaine de tests. Si les auteurs décrivent des données classiques, recueillies de façon classique pour réaliser une comparaison qui semblerait classique, et qu'ils utilisent des tests statistiques aux noms imprononçables, qui n'existent pas dans les livres statistiques de base, le lecteur doit se méfier.

Enfin, les données doivent avoir été analysées en suivant le protocole originel. Notamment, chercher dans les données des résultats intéressants a posteriori, par exemple par des analyses en sous groupes peut entraîner des conclusions faussement positives. Ces analyses ne peuvent pas entraîner des conclusions majeures, mais doivent simplement être génératrices d'hypothèses nouvelles, qui devront être testées par d'autres études.

#### 4) Tests unilatéraux (One Tailed Test) ou bilatéraux (Two Tailed Test) ?

Les tests bilatéraux s'intéressent aux différences entre les deux groupes d'un côté ou de l'autre, c'est-à-dire si un des deux groupes, peu importe lequel, est supérieur à l'autre. Ils doivent être privilégiés. Alors que les tests unilatéraux s'intéressent à un seul côté de la différence, c'est-à-dire un groupe spécifique supérieur à l'autre groupe. Les auteurs, s'ils ont utilisé des tests unilatéraux doivent expliquer pourquoi et donner clairement l'hypothèse statistique testée, à savoir quel groupe était considéré supérieur à l'autre. Des tests unilatéraux n'ont aucune puissance statistique pour montrer l'inverse de l'hypothèse testée a priori.

#### 5) Valeurs aberrantes :

Des résultats bizarres ou tout au moins non attendus peuvent refléter des particularités chez un ou plusieurs sujets, des erreurs de mesure, des erreurs d'interprétation de la lecture des données d'un appareil, des erreurs de calcul. Lorsqu'un résultat est très éloigné de ce que l'on aurait pu attendre, il est possible que les auteurs l'excluent de l'analyse. Cela ne doit se faire bien entendu qu'avec prudence, en s'étant assuré qu'il ne s'agit pas d'un résultat réel, traduisant des particularités chez un sujet donné dans le groupe. La correction statistique pour des valeurs aberrantes est possible mais requiert des méthodes sophistiquées.

#### 6) Est-ce que les valeurs de « p » ont été calculées et surtout interprétées de façon appropriées ?

La valeur de « p » correspond à la probabilité que le résultat constaté soit lié au hasard. Dans la pratique scientifique standard, de façon totalement arbitraire, lorsque le « p » est inférieur ou égal à 5%, on considère qu'il existe une signification statistique et que le résultat n'est pas lié au hasard. Cela veut donc dire par définition qu'une association significative peut être retrouvée dans un cas sur vingt par hasard.

Une des conséquences est de se méfier de la réalisation de nombreux tests statistiques à partir du même jeu de données. En effet, dans ce cas, il est possible d'obtenir une association simplement liée au hasard dans cinq cas sur cent. Il faut donc prendre cela en compte et effectuer des corrections : habituellement il s'agit de la correction de Bonferroni qui requiert simplement de diviser le seuil de 5% par le nombre de tests réalisés et de considérer les valeurs de « p » inférieures à ce nouveau seuil calculé comme statistiquement significatives (exemple : pour 10 tests, nouveau seuil à 0,005).

A l'inverse, une valeur de « p » supérieure à 0,05 peut avoir deux explications : soit il n'existe effectivement pas de réelle différence entre les groupes, soit il y avait trop peu de sujets

pour la mettre en évidence (manque de puissance). **C'est pourquoi le calcul du nombre de sujets nécessaires est extrêmement important.** Si ce calcul avait été correctement réalisé et qu'il a été atteint, un test non significatif ne résulte pas d'un manque de puissance. On peut alors affirmer que la différence utilisée pour calculer le nombre de sujets nécessaires n'existe pas, au risque bêta consenti (habituellement 20% ou 10%).

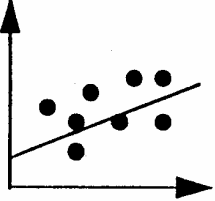
Enfin, une valeur de « p » donnée seule ne suffit pas. Il faut donner également les valeurs comparées. Il est de bonne pratique de donner également les intervalles de confiance des paramètres estimés. Cet intervalle de confiance correspond à la variation du paramètre dans une population plus large dont serait issu l'échantillon de l'étude. Ces intervalles de confiance peuvent être donnés, que l'étude montre une association statistique ou non. Ces intervalles de confiance peuvent être calculés pour l'ensemble des paramètres intéressants dans une étude, que ce soit un pourcentage, une moyenne ou d'autres paramètres tels que sensibilité, spécificité, etc ... L'intervalle de confiance permet de se rendre compte si les résultats de l'étude sont suffisamment précis. Si la variation du paramètre est trop importante, même si l'étude a un résultat statistiquement significatif, il est prudent de se méfier. L'interprétation de l'intervalle de confiance doit être réalisé grâce au bon sens clinique.

## **CONCLUSION :**

Ces quelques pages n'ont bien entendu pas pour but de faire apprendre la biostatistique, même basique, mais simplement de donner quelques clés de lecture d'articles incluant de telles analyses. L'exercice le plus difficile sera sans doute de savoir si le test réalisé était le plus approprié. Cela nécessite de la pratique. Il faut revenir aux deux questions de base : Quelles variables sont-elles comparées ? De quels types étaient ces variables ? C'est-à-dire qualitatives (à deux ou plus de deux classes) ou quantitatives.

## Tableau résumé des différents tests utilisés en Biostatistique

Cas	Tests paramétriques	Conditions	Tests non-paramétriques	Conditions
<b><u>Liaison de 2 variables qualitatives :</u></b> (comparaison de 2 %)				
<ul style="list-style-type: none"> <li>• <i>Indépendantes</i></li> </ul>	$\varepsilon \text{ ou } t = \frac{(P_a - P_b)}{\sqrt{\frac{pq}{n_a} + \frac{pq}{n_b}}}$	avec $p = \frac{n_a p_a + n_b p_b}{n_a + n_b}$ et $q = 1 - p$  $nP_a, nP_b, nQ_a, nQ_b \geq 5$	<b>Test du Chi<sup>2</sup></b>  $\chi^2 = \frac{\sum (O_i - C_i)^2}{C_i}$	Si $C_i \geq 5$  $ddl = (L - 1)(C - 1)$ Si $L$ ou $C = 1$ , $ddl = L - 1$ ou $C - 1$
<ul style="list-style-type: none"> <li>• <i>Appariées</i></li> </ul>	$\varepsilon = \frac{a - b}{\sqrt{a + b}}$	$a + b \geq 10$ $a$ et $b$ : paires discordantes	<b>Test du Chi<sup>2</sup> corrigé de Yates</b> si $C_i \geq 3$ ou $< 5$ Sinon, <b>test exact de Fisher</b>  <b>Test du chi<sup>2</sup> de Mac Némar</b>	$\chi_{MN}^2 = \frac{(a - b)^2}{(a + b)}$ Si $a + b \geq 10$ ( $ddl = 1$ )
<b><u>Liaison entre 1 variable qualitative à 2 classes et 1 variable quantitative</u></b> (comparaison de 2 moyennes) :				
<ul style="list-style-type: none"> <li>• <i>2 échantillons indépendants</i></li> </ul>	$\varepsilon = \frac{ m_a - m_b }{\sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}}}$	$n_a$ et $n_b \geq 30$	<b>Test de Mann-Whitney</b>	
	$t = \frac{ m_a - m_b }{\sqrt{s^2 \left( \frac{1}{n_a} + \frac{1}{n_b} \right)}}$	$s^2 = \frac{n_a s_a^2 + n_b s_b^2}{n_a + n_b - 2}$		
		$ddl = n_a + n_b - 2$ ; $n_a$ ou $n_b \leq 30$ Loi normale ; égalité des variances testée par $F = \frac{s_a^2}{s_b^2}$ ( $s_a^2 > s_b^2$ )		
<ul style="list-style-type: none"> <li>• <i>2 échantillons appariés</i></li> </ul>	$\varepsilon = \frac{\bar{d}}{\frac{s_d}{\sqrt{n - 1}}}$	$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n - 1}}}$	<b>Test de Wilcoxon sur séries appariées</b>	
	Si $n \geq 30$	Si $n < 30$	distribution normale des $d_i$ $ddl = n_d - 1$	

Cas	Tests paramétriques	Conditions	Tests non-paramétriques	Conditions
<b><u>Liaison de 1 variable qualitative à n &gt; 2 classes</u></b>				
<b><u>à 1 variable quantitative</u></b>				
(comparaison de n (> 2) moyennes)				
• <i>Indépendantes</i>	A.N.O.V.A.	Variances égales (test de Bartlett) Loi normale	Test de Kruskal Wallis	
• <i>Appariées</i>	A.N.O.V.A.	“ “ “	Test de Friedman	
	1 facteur contrôlé 1 facteur aléatoire			
<b><u>Liaison entre 2 variables quantitatives</u></b>				
	Coefficient de corrélation (r) distribution liée normale et variance constante ; ddl = n - 2		Test de Spearman $r_s$ ( $\rho$ )	
	$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \text{ et } t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2}$		$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$	
	Régression si 1 variable est contrôlée			
	Modèle linéaire : $y = px + q$			
	$p = r \frac{s_y}{s_x}$			
	$q = \bar{y} - p\bar{x}$			
<b><u>Comparaison de 2 ou n taux de survie</u></b>			Log-Rank	Pas de croisement net des courbes de survie

**CHAPITRE VI :**  
**RECOMMANDATIONS SYNTHETIQUES**  
**POUR LA REALISATION**  
**D'UNE ANALYSE DE LECTURE CRITIQUE D'ARTICLE**

**Recommandation n°1 :**

Savoir définir le type de l'étude, essai clinique, épidémiologique, article diagnostique en précisant à chaque fois les caractéristiques de cette étude. En règle générale, cette information se trouve à la fin de l'introduction de l'article.

**Recommandation n°2 :**

Savoir rechercher les critères de jugement. Un seul critère principal devrait être présenté par les auteurs, mais il peut exister de nombreux critères secondaires. Les critères de jugement se trouvent en général dans la section « matériels et méthodes » dans les méthodes de l'article. Il faut être capable également de discerner si ces critères sont bien adaptés à la question posée.

**Recommandation n°3 :**

Savoir analyser et décrire la population sélectionnée par les auteurs. Ceux-ci doivent avoir correctement défini les patients, mais également la pathologie c'est-à-dire les cas avec une description soignée de ce qui a permis d'inclure (critères d'inclusion) ou de ne pas inclure (critères de non inclusion) les patients dans l'étude.

**Recommandation n°4 :**

Vérifier que le nombre de sujets nécessaires a été estimé a priori, quel que soit le type d'étude.

**Recommandation n°5 :**

Savoir analyser la présentation des résultats. Le texte n'est pas forcément long mais les tableaux viennent en complément du texte. Les résultats principaux doivent être énoncés, notamment ceux concernant les critères sélectionnés par les auteurs. La présentation des résultats descriptifs doit être adéquate, en fonction du type de variables analysées. Les tests statistiques doivent être appropriés.

**Recommandation n°6 :**

S'il s'agit d'un essai thérapeutique, savoir chercher les 3 points critiques : randomisation, conduite en double insu, analyse en intention de traiter.

***Recommandation n°7:***

Savoir analyser la discussion de l'article. Les premières lignes de la discussion reprennent le résultat principal de l'étude. Cette discussion est en général structurée en trois parties. Après un bref rappel de leurs résultats, les auteurs doivent donner éventuellement les limites. On y trouve d'ailleurs souvent une discussion des biais par les auteurs eux-mêmes. On doit ensuite trouver une discussion par rapport aux résultats de la littérature et la dernière partie replace l'étude par rapport aux résultats généraux. Il ne peut pas y avoir dans la discussion de données qui n'ont pas été présentées dans les résultats.

***Recommandation n° 8:***

Vérifier que les références sont conformes et notamment celles qui ont permis de présenter la méthode utilisée.