

FACULTE DE MEDECINE DE LIMOGES

---

**LA LITTERATURE SCIENTIFIQUE :**

**LA RECHERCHE BIBLIOGRAPHIQUE**

**LA LECTURE CRITIQUE**

**LA REDACTION D'UN RESUME D'ARTICLES SCIENTIFIQUES**

**LES STATISTIQUES DANS LA LITTERATURE SCIENTIFIQUE**

Professeur A. VERGNENEGRE

Professeur P. M. PREUX

Année 2002 - 2003

Les progrès médico-scientifiques sont source d'une masse considérable d'informations publiées dans la littérature.

La lecture de ces informations peut se faire dans plusieurs cadres :

- par le clinicien qui veut se tenir au courant des évolutions susceptibles d'influencer sa pratique,
- par des recherches visant à évaluer de nouvelles technologies, dans le cadre d'un travail scientifique de recherche (thèse de doctorat en médecine par exemple).

Le cadre général de cette littérature pour un médecin est donc le suivant :

- se tenir au courant,
- comprendre (la physio-pathologie, l'épidémiologie, ...),
- changer (les attitudes diagnostiques, préventives et thérapeutiques),
- s'informer,
- comprendre, admettre ou contester les contraintes économiques.

Cette lecture se heurte dans tous les cas à l'écueil que constitue la quantité de données publiées : plus de 30 000 articles sont recensés chaque mois dans la banque de données Medline. Ainsi, celui qui voudrait mettre à jour ses connaissances dans la maladie d'Alzheimer par exemple, se trouverait confronter à une liste de plus de 23 000 articles sur trois ans.

Cet état de chose rend indispensable une démarche en plusieurs étapes que nous allons décrire :

- rechercher l'information,
- sélectionner et analyser l'information,
- synthétiser l'information dans un résumé,
- savoir critiquer la méthodologie statistique.

**CHAPITRE I :**  
**IDENTIFICATION DES INFORMATIONS :**  
**LA RECHERCHE BIBLIOGRAPHIQUE**  
**OU RECHERCHE DOCUMENTAIRE**

C'est une démarche systématique et explicite d'identification, de récupération et de traitement bibliographique de données publiées ou non. Elle a pour but l'identification d'informations pertinentes et doit tendre vers l'exhaustivité.

## **I- METHODOLOGIE DE LA STRATEGIE DE LA RECHERCHE DOCUMENTAIRE :**

Elle est facilitée par la connaissance des bases de données utilisées et des mots clés.

### *A- L'étape préalable :*

Cette étape permet de bien analyser un projet avant de lancer une recherche documentaire approfondie. Il faut définir, d'une part l'étendue du projet et son domaine : est-ce un domaine où la littérature est très abondante ? est-ce un domaine où, au contraire, la littérature est pauvre et faite d'études peu adaptées ? Une fois le thème bien défini, il est possible de choisir quelques critères soit d'exclusion, soit d'inclusion à respecter lors de la stratégie de la recherche :

- l'un des critères d'exclusion le plus usité est la langue. Seules les publications de langue anglaise, française sont souvent recherchés pour des raisons simples de compréhension.
- une limitation sur les années de recherche est généralement effectuée. La période retenue est variable selon les thèmes.

### *B- Elaboration des stratégies de recherche documentaire :*

- identification du thème,
- recherche approfondie à partir d'une première interrogation,
- interrogation systématique de plusieurs banques de données,
- recherche sur des banques de données spécifiques,
- recherche de « la littérature grise ». La littérature grise est une littérature non indexée dans les catalogues officiels d'édition et les circuits conventionnels. Il s'agit souvent d'un rapport de documents qui ne sont jamais publiés. Cette recherche se fait par des contacts directs auprès des Sociétés Savantes, par la consultation d'Internet, par la recherche d'ouvrages.
- dépouillement des sommaires de revues.

## **II- LES SOURCES D'INFORMATION :**

### *A- Les banques de données bibliographiques :*

#### 1) les banques de données médicales :

Medline : c'est la base créée en 1966 par la National Library of Medicine (à Bethesda – Etats-Unis). C'est la plus connue et la plus utilisée. Elle contient plus de 7 millions de citations dont plus de 5 millions concernent l'être humain. Elle indexe la littérature scientifique de plus de 4 000 revues internationales. Sa couverture est essentiellement anglo-saxonne. Des termes normalisés et structurés (mots-clés) sont utilisés pour indexer les documents et les questions. L'ensemble de ces termes constitue un thesaurus. Le thesaurus de Medline s'appelle MeSH (Medical Subject Headings). Quand un article est indexé dans Medline, 10 à 12 termes du MeSH y sont attachés. Le MeSH est régulièrement actualisé.

Enbase : c'est une banque de données médicales produite aux Pays-Bas. Elle indexe environ 3 700 journaux. Elle possède deux particularités : une orientation sur la pharmacologie et la toxicologie tout en traitant les autres aspects de la médecine et une bonne couverture de la littérature européenne.

Pascal : c'est une banque de données française produit par l'Institut National de l'Information Scientifique et Technique (INIST). Créée en 1973, cette base multidisciplinaire concerne les domaines aussi divers que la biologie, la médecine, les disciplines fondamentales de la physique et de la chimie, les sciences de l'ingénieur, l'océanographie. La littérature française représente environ 10% des 12 millions de références bibliographiques.

Biosis : produite aux Etats-Unis, elle contient plus de 11 millions de citations. Elle est particulièrement intéressante pour la recherche de résumés de congrès.

De nombreuses autres bases de données sont plus spécifiques, telles que les bases de données sur le SIDA (Aidline), le cancer (Cancerlit), Cinahl pour les infirmiers et les paramédicaux, Helmis pour la gestion des services de santé, Psyclit pour le domaine de la psychiatrie et les Current Contents pour le sommaire des journaux.

#### 2) La Cochrane Library :

C'est une banque de données électronique, disponible sous forme de disquettes ou de CD-Rom. Elle a été initiée en 1995 en Grande-Bretagne. Elle est constituée de plusieurs

banques. La base des revues systématiques qui met à la disposition des synthèses sur des thèmes de la littérature régulièrement actualisés. Une base d'abstracts de revues d'efficacité qui recense les revues de la littérature de bonne qualité méthodologique publiées dans des revues médicales et des résultats d'Agence Internationale d'Evaluation Technologique.

3) les banques de données économiques et industrielles :

Elles permettent d'obtenir des informations provenant de revues internationales, nationales et locales, des rapports financiers, des rapports d'activité. Elles peuvent avoir une ouverture internationale ou européenne.

DELPHES est produite par la Chambre de Commerce et d'Industrie de Paris. Elle donne des informations bibliographiques provenant d'environ 900 journaux.

*B- La littérature grise :*

1) les banques de données de littérature grise :

La banque de données de Santé Publique a été créée en France en 1994 et contient plus de 150 000 références.

HEALTHSTAR contient de la littérature scientifique émanant de revues, de rapports techniques, de résumés de congrès, de chapitres de livres et de documents gouvernementaux. Elle couvre plus particulièrement les champs de l'évaluation technologique et la recherche en administration de santé.

NHS Economic Data Base contient des informations sur les évaluations économiques.

2) les agences et institutions gouvernementales, les Sociétés Savantes :

Ces associations professionnelles tiennent à jour des rapports qui peuvent être accessibles sur demande.

*C- Dépouillement systématique des sommaires de revues :*

Les principales revues à lire sont le Lancet, le New England Journal of Medicine, la Presse Médicale, le British Medical Journal, The Annals of Internal Medicine, JAMA, la Revue du Praticien.

### **III- LES CRITERES D'INTERROGATION :**

Comme nous l'avons souligné, ils seront indépendants de la base interrogée. Ils peuvent concerner :

- les périodes de temps,
- la langue,
- les mots-clés,
- les thèmes (études in vitro, études sur l'animal, études sur l'être humain),
- les types d'articles (abstracts, revues de la littérature).

### **IV- LE CHOIX DES MOTS CLES :**

Lorsque l'on interroge dans la base Medline, c'est le MeSH qui est utilisé. Il concerne essentiellement les mots qui vont décrire en résumé, principalement les articles. Ils peuvent être complétés par des Subheading qui sont des restricteurs (traitement, diagnostic, pronostic, dépistage).

Si l'on prend l'exemple de l'épidémiologie du cancer du poumon, il faut d'abord rentrer le terme cancer du poumon qui donnera sur une ou deux années 24 000 références, puis épidémiologie qui donnera également 15 000 à 20 000 références, et c'est le croisement de ces deux mots, par la syntaxe « et » qui permettra d'obtenir les références souhaitées (900 à 1 000).

### **V- LA RECHERCHE EN PRATIQUE :**

Comme nous l'avons vu, il faut d'emblée s'orienter vers des bases de données existantes et le Medline en est la principale. Les Universités ont mis en place une base OVID qui permet, à partir du Medline, de faire des recherches par sujet, auteur, journal et des recherches combinées. Cette base est accessible par Internet et permet des interrogations simples et rapides. Il est possible dans cette base d'ajouter des subdivisions avec des abréviations, par exemple l'éducation de l'asthme se recherche par asthma/ed. Sur la recherche par Internet, il existe au niveau de OVID des guides qui permettent de comprendre comment se fait cette recherche, ainsi que les différentes interrogations.

## **VI- EN CONCLUSION :**

La recherche bibliographique nécessite une certaine pratique. Elle est cependant indispensable dès que l'on va avoir à réaliser un travail, que ce soit un exposé dans un service ou pour un certificat et bien sûr son travail de thèse de médecine.

Il ne faut souvent ne pas se contenter du Medline.

Pour l'ANAES, lors de la recherche concernant des recommandations de pratique clinique, la méthodologie suivante est utilisée :

- Medline
- Healthstar,
- Enbase,
- Pascal,
- Cochrane Data Base,
- ± littérature grise.

**CHAPITRE II :**

**LA LECTURE CRITIQUE DES ARTICLES SCIENTIFIQUES :**

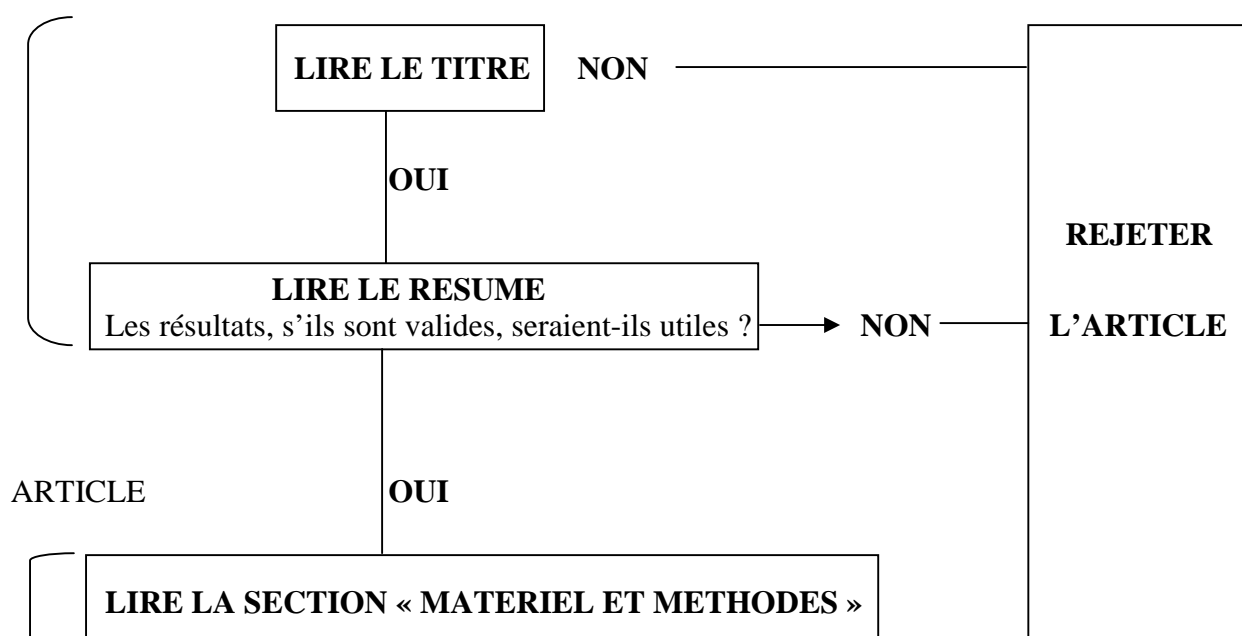
**SELECTION DE LA LITTERATURE**

## **I- LA SELECTION DES ARTICLES :**

Cette étape doit tenir compte de plusieurs facteurs :

- la présence ou non d'un résumé (environ 75% des articles en ont un),
- l'accessibilité et/ou le coût des documents,
- le temps alloué à cette recherche,
- la qualité de la revue qui a publié l'article.

De façon assez sommaire, on peut schématiser le principe sur la figure ci-dessous :



La première sélection est effectuée rapidement sur le titre et la revue. Une revue difficile à se procurer ou confidentielle peut être écartée (par exemple : le journal de médecine interne d'une province du sud de la Nouvelle Zélande), ainsi que le titre, si celui-ci n'est pas suffisamment évocateur de l'intérêt pour l'individu qui réalise la recherche. Lorsqu'un résumé est disponible, certains critères peuvent être définis qui permettent rapidement de sélectionner ou non l'article.

## **II- L'ANALYSE DE LA QUALITE DE LA LITTERATURE :**

Un certain nombre de questions préliminaires doivent être posées.

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

Objectif et hypothèse de travail

⇒ *Quelle est le type de l'étude ?*

- expérimentales (animaux, volontaires sains)
- essais cliniques (souvent thérapeutiques et comparatifs) ?
- enquêtes (touchant une population ou un échantillon),
- 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 quelle est alors son apport par rapport aux travaux antérieurs.

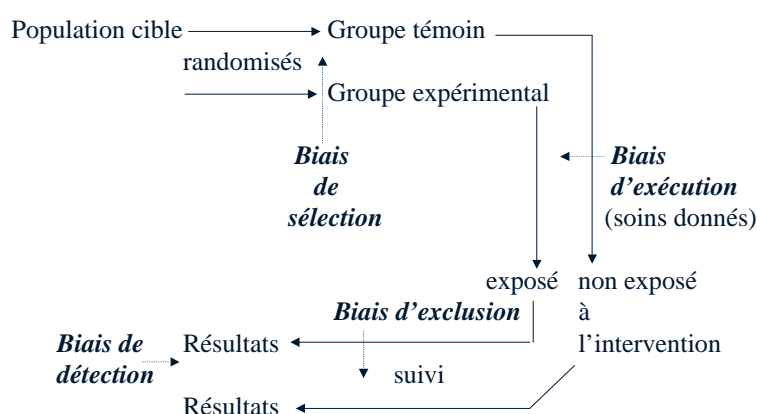
- qui sont les sujets de l'étude (recrutement, inclusion, non-inclusion, perdu de vue).

⇒ *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 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

⇒ *Quel est le niveau de preuve de l'article ?*

On distingue plusieurs niveaux de preuve dans un article scientifique, essentiellement cinq niveaux. Les deux tableaux ci-dessous montrent la classification américaine et canadienne des niveaux de preuve.

---

### **Niveau de preuve scientifique apporté par une étude selon la force de protocole**

*(adapté du score utilisé par l'American College of Chest Physicians)*

---

Niveau 1	Essais contrôlés randomisés avec résultats indiscutables (méthodologiquement)
Niveau 2	Essais contrôlés non randomisés bien conduits
Niveau 3	Essais prospectifs non contrôlés bien menés (suivi de cohorte par exemple)
Niveau 4	Etudes cas-témoins : essais contrôlés présentant des biais
Niveau 5	Etudes prospectives et cas cliniques (série de malades) Toute étude fortement biaisée

---



---

### **Niveau d'efficacité d'une intervention médicale**

*(adapté d'après le Canadian Task Force on the Periodic Health)*

---

I	Preuves obtenues par au moins un essai comparatif randomisé
II 1	Preuves obtenues au moyen d'essais comparatifs non randomisés
II 2	Preuves obtenues par des études de cohorte ou des cas-témoins, de préférence multicentriques
II 3	Preuves obtenues par des comparaisons de séries non contemporaines
III	Avis d'experts Etudes descriptives

---

Ces niveaux de preuve doivent être formalisés, basés sur des grilles de lecture et caractérisent la capacité de l'étude à répondre à la question posée. Ils font référence à l'évidence scientifique qui résulte de la lecture de tous les articles ou sont basés éventuellement sur un accord professionnel lorsqu'il s'agit uniquement de l'opinion des professionnels de santé.

### **III- L'ANALYSE DE LA QUALITE DE LA LITTERATURE :**

Après les deux premières phases d'identification et de sélection de l'article, l'analyse de la qualité va permettre de retenir ceux dont les résultats vont être utilisés pour le travail recherché. Cette analyse constitue l'étape initiale pour préparer des recommandations de pratique et des références professionnelles. Il est possible, éventuellement, de se servir d'une grille d'analyse pour tout type de publication. Ce type de grille comporte souvent 8 critères d'évaluation :

- l'objectif de l'étude,
- le type du protocole,
- le facteur étudié,
- les critères de jugement,
- la population source,
- les facteurs de confusion et les biais,
- les analyses statistiques et les conclusions des auteurs.

Le tableau ci-après reprend l'analyse d'une publication médicale originale.

En fait, il est souvent préférable de séparer les analyses en fonction du type d'article.

<b>ANALYSE D'UNE PUBLICATION MEDICALE ORIGINALE</b>
---

La question est-elle clairement posée ?	Oui	Non
Les bases scientifiques sur lesquelles repose l'étude sont-elles citées ?	Oui	Non
De quel type d'étude s'agit-il ?	Expérimentale Randomisée Contrôlée non randomisée Autre	Observation Cas-témoin Prospective Rétrospective, transversale, rapport d'observations, opinions d'experts, autres
Tous les termes techniques employés sont-ils bien définis ?	Oui	Non
Les variables et les résultats finaux analysés sont-ils clairement identifiés ?	Oui	Non
Quels sont les résultats finaux analysés (efficacité, effets secondaires) :		
La mesure et le choix des résultats finaux sont-ils corrects ?	Oui	Non
Le choix des variables et leur méthode de mesure sont-ils corrects ?	Oui	Non
Quels sont les biais de mesure identifiables ?	Oui	Non
Quelle est la population étudiée ?		
Quelle est la population témoin ?		
Comment les échantillons ont-ils été constitués ?		
La constitution des populations étudiées est-elle correcte ?	Oui	Non
La taille des échantillons est-elle adaptée ?	Oui	Non
Les méthodes statistiques sont-elles exposées en détail ?	Oui	Non
Les méthodes statistiques sont-elles appropriées ?	Oui	Non
La durée de suivi est-elle suffisante ?	Oui	Non
Les facteurs confondant potentiels sont-ils pris en compte ?	Oui	Non
Les résultats sont-ils présentés de façon claire, détaillée, objective ?	Oui	Non
Les résultats sont-ils cohérents ?	Oui	Non
Les conclusions sont-elles valides ?	Oui	Non

*Toutes les réponses du type «oui-non» peuvent être remplacées par une cotation de 0 (non) à 1 (oui).  
Expérimentale = 5, cas témoin = 4, contrôlée prospective = 3, autre = 2.*

*A- Analyse d'une revue de synthèse :*

Une revue de synthèse présente l'analyse de travaux originaux déjà publiés. Si cette synthèse a suivi une méthode explicite et valide, elle peut être prise comme document de référence, ce qui évite ainsi une analyse détaillée de la littérature antérieure à ce document.

1) Les objectifs sont-ils clairement exposés ?

L'hypothèse attestée est expliquée. La question à laquelle l'auteur essaye de répondre est correctement exposée et segmentée en sous questions.

2) La méthodologie :

Les procédures de sélection de la littérature doivent être décrites.

Les sources de données doivent être décrites. La précision des procédures utilisées pour obtenir les informations doit être donnée.

Les critères d'inclusion et d'exclusion et le mode de sélection des articles doivent être énumérés.

Les méthodes d'analyse doivent être précisés : caractéristique des lecteurs, anonymisation des articles, lecture en double-aveugle, ...

La méthode utilisée pour analyser la synthèse de travaux doit être présentée.

Les résultats doivent être correctement explicités et la validité des arguments sur lesquels l'auteur fonde ses conclusions discutés.

<b>GRILLE DE LECTURE DES REVUES DE SYNTHÈSE</b>
---

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

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

	Totalem	Partiellem	Pas du tout
1. Les objectifs de la revue de synthèse sont clairement exposés	q	q	q
2. Méthodologie			
2.1. <i>Procédures de sélection</i>			
• L'auteur décrit ses sources de données	q	q	q
• Les critères de sélection des études sont pertinents	q	q	q
• Les critères d'inclusion et d'exclusion des articles sont décrits	q	q	q
• Les études non publiées sont prises en compte	q	q	q
2.2. <i>Méthode d'analyse</i>			
• Les modalités de la lecture critique sont précisés (lecteurs, grille de lecture...)	q	q	q
• L'auteur présente la méthode utilisée pour réaliser la synthèse des résultats.	q	q	q
3. Résultats			
• L'auteur décrit les résultats	q	q	q
• L'auteur commente la validité des études choisies	q	q	q
• Ses conclusions s'appuient sur des données fiables dont les sources sont citées	q	q	q
4. Applicabilité clinique			
• La revue de synthèse permet de répondre en pratique à la question posée	q	q	q

**Commentaires :**

*B- 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 IV). 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 :**

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

**Commentaires :**

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

Le 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 :**

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

**Commentaires :**

*D- Analyse d'un article de causalité :*

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</b>
--

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

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

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

**Commentaires :**

*E- Analyse d'un article de pronostic (analyse 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 DE PRONOSTIC (ANALYSE 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	q	q	q
2. Méthodologie	q	q	q
• Les modalités de constitution de la cohorte sont précisées.	q	q	q
• Tous les patients de la cohorte ont été identifiés au même stade de la maladie.	q	q	q
• Les critères d'inclusion et d'exclusion sont spécifiés et adéquats.	q	q	q
• Les biais possibles sont exposés et les méthodes pour les prendre en compte sont décrites.	q	q	q
• Le suivi est complet et correctement réalisé.	q	q	q
• Les critères de jugement sont pertinents, fiables et tous utilisés.	q	q	q
3. Les résultats			
• L'interprétation de ces critères est objective.	q	q	q
• Les résultats sont ajustés pour les autres facteurs pronostiques.	q	q	q

**Commentaires :**

*F- Analyse d'un article épidémiologique:*

La grille de lecture est décrite ci-après et prend en compte essentiellement la méthode et les résultats.

<b>GRILLE DE LECTURE D'UN ARTICLE EPIDEMIOLOGIQUE</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 de l'étude sont clairement définis	q	q	q
2. Méthodologie	q	q	q
• Les caractéristiques de la population sont décrits	q	q	q
• Les critères d'inclusion et d'exclusion sont précisés et adéquats	q	q	q
• Les qualités et les modalités de recueil des données sont précisés.	q	q	q
3. Analyse des résultats			
• L'analyse statistique est adaptée	q	q	q
• Les facteurs de confusion et les biais sont pris en compte.	q	q	q
• Les résultats sont vérifiables à partir des données brutes.	q	q	q

**Commentaires :**

*G- Analyse d'un document de recommandations :*

Celui-ci a une certaine importance car il faut être capable également de critiquer certaines recommandations de pratique que nous décrirons dans un chapitre spécifique. Il faut pouvoir préciser le contexte d'élaboration, l'objectif, les populations concernées. La méthode des recommandations doit être décrite, ainsi que leur analyse.

La grille de lecture est décrite ci-après.

<b>GRILLE DE LECTURE D'UN DOCUMENT DE RECOMMANDATIONS</b>
---

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

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

**Promoteur :**

**1. Contexte et objectifs**

	OUI	NON	Partiellement
• Le contexte d'élaboration des recommandations est précisé	q	q	q
• L'objectif des recommandations est précisé	q	q	q
• Les populations concernées par les recommandations sont précisées	q	q	q

**2. Méthodologie**

• La méthodologie employée pour l'élaboration des recommandations est clairement présentée.	q	q	q
• Les critères de jugement des études qui ont servi à élaborer les recommandations, sont explicités.	q	q	q
• L'argumentaire des recommandations est précisé	q	q	q

**3. Les recommandations**

• Les conclusions et recommandations correspondent aux informations analysées.	q	q	q
• Les recommandations sont claires et précises.	q	q	q
• Les recommandations sont adaptées à la pratique clinique quotidienne et aux cibles.	q	q	q

4. Un processus de validation est mentionné.	q	q	q
--	---	---	---

**Commentaires :**

#### *H- Sélection et analyse de la littérature économique :*

La sélection des études économiques : les critères d'inclusion d'un article économique doivent être les suivants :

- la perspective (point de vue), à partir de laquelle l'évaluation a été menée,
- le bénéfice net individuel apporté pour le patient, par la technique ou la pratique médicale,
- l'évaluation doit être comparative, deux alternatives au moins, thérapeutique et diagnostique sont soumises au calcul. Les auteurs doivent spécifier quels types de coûts ils ont utilisé, le taux d'actualisation lorsque l'étude s'est déroulée sur plusieurs années,
- des analyses de sensibilités doivent être pratiquées.

La grille d'analyse d'un article économique est décrit ci-après. Elle prend en compte la définition du cadre de l'évaluation, l'identification de la mesure et la valorisation des coûts, la description de l'analyse comparative et la prise en compte de l'incertitude.

<b>GRILLE D'ANALYSE D'UN ARTICLE ÉCONOMIQUE</b>
---

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

Rev/Année/Vol/Pages \_\_\_\_\_

**Thème de l'article :**

### A. CADRE DE L'ÉVALUATION

#### 1) QUELLE PERSPECTIVE EST ADOPTÉE DANS L'ÉTUDE ?

	OUI	NON
• La société	r	r
• Le patient	r	r
• L'assurance maladie	r	r
• Les assurances complémentaires	r	r
• L'établissement de soins	r	r

#### 2) QUELLE EST LA NATURE DE LA COMPARAISON ?

*A préciser*

### B. QUANTIFICATION/MESURE DES COÛTS : QUELS SONT LES COÛTS RETENUS ?

#### 1) COÛTS DIRECTS MÉDICAUX ET NON MÉDICAUX

	OUI	NON	<b>Si oui, mode de valorisation</b>		
			Coût réel	Coût standard	Coût nominal
• Coût des stratégies valorisées	r	r	r	r	r
<i>A préciser</i>					
Avec prise en compte :					
- du personnel	r	r			
- des consommables utilisés	r	r			
- de l'amortissement du matériel	r	r			
- de sa maintenance	r	r			
• Coût des effets secondaires liés aux stratégies	r	r	r	r	r

#### 2) COÛTS INDIRECTS

	OUI	NON	<b>Si oui, mode de valorisation</b>
• Coûts indirects pour le patient	r	r	Coût des pertes de production r

#### 3) COÛTS INTANGIBLES

	OUI	NON
• Coûts de l'anxiété	r	r
• Coûts de la douleur	r	r

### C. ÉVALUATION COMPARATIVE

	OUI	NON
1) ÉTUDE DE MINIMISATION DES COÛTS	r	r
2) ÉTUDE COUT - EFFICACITE		
• Critères d'efficacité :		
- années de vie sauvées	r	r
- autres	r	r

**3) ÉTUDE COUT – UTILITE**

- QALY	r	r
- autres	r	r

**4) ÉTUDE COUT – BENEFICE**

- coûts évités	r	r
- disposition à payer	r	r
- autres	r	r

**D. PRISE EN COMPTE DU TEMPS**

	OUI	NON
• Actualisation des coûts	r	r
• Actualisation des bénéfices	r	r

Cette grille est adaptée selon les thèmes traités. Certaines questions sont donc à développer selon les sujets.

#### **IV- NOTION DE NIVEAUX DE PREUVE ET GRADUATION DES RECOMMANDATIONS :**

A l'issue de cette analyse, et conformément à ce qui a été décrit, la décision d'utiliser l'article doit pouvoir être prise.

Les articles analysés le sont au sein d'un tableau de synthèse faisant apparaître :

- la référence de la publication,
- le protocole,
- le facteur,
- les caractéristiques de la population de l'étude,
- les critères de jugement,
- les résultats,
- éventuellement les conclusions de l'auteur.

##### *A- Le niveau de preuve :*

Nous avons déjà détaillé le niveau de preuve de ces articles au début de ce chapitre.

##### 1) Niveau de preuve d'une étude :

Le niveau de preuve d'une étude caractérise la capacité d'étude à répondre à la question posée. Cette capacité se juge par la correspondance de l'étude au cadre du travail réalisé, d'autre part par les caractéristiques suivantes :

- l'adéquation du protocole à la question posée,
- l'existence ou non de biais importants dans la réalisation,
- la puissance de l'étude et la taille de l'échantillon. La classification générale peut être proposée à partir des classifications de la littérature et des composants vus ci-dessus.
  - un fort niveau de preuve correspond à une étude dont le protocole est adapté pour répondre mieux à la question posée, la réalisation est effectuée sans biais majeur, l'analyse statistique est adaptée aux objectifs, la puissance est suffisante,
  - un niveau intermédiaire est donné par une étude de protocole similaire, présentant une puissance nettement insuffisante des anomalies mineures,
  - un faible niveau de preuve peut être attribué aux autres types d'études.

##### 2) L'évidence scientifique :

Elle est appréciée lors de la synthèse des résultats de l'ensemble des études sélectionnés. La gradation de l'évidence scientifique s'appuie sur :

- l'existence de données de la littérature,

- des niveaux de preuve,
- la cohérence de leurs résultats.

Elle constitue la conclusion des tableaux de synthèse.

*B- La notion d'accord professionnel :*

Un accord professionnel exprime une opinion quasi-unanime. Selon cette opinion, une méthode de dépistage, des diagnostics, de prévention ou de traitement, peut être proposée ou non dans une indication donnée. Cet accord professionnel correspond à un fait d'observation.

Un accord professionnel nécessite d'être validé par un groupe de lecture externe au groupe de travail à l'origine de l'accord.

Ce groupe d'experts doit être représentatif des différents professionnels de santé concernés et de différentes modalités d'exercice.

*C- Grade des recommandations :*

La graduation des recommandations selon l'ANAES se fait en trois catégories : A – B – C (cf tableau suivant)

<b>Niveau de preuve scientifique de la littérature et force des recommandations</b> <i>(adapté du score de Sackett)</i>
--

NIVEAU DE PREUVE SCIENTIFIQUE FOURNI PAR LA LITTERATURE	GRADE DES RECOMMANDATIONS
<b>Niveau 1</b>  - Essais comparatifs randomisés de forte puissance - Méta-analyse d'essais comparatifs randomisés - Analyse de décision basée sur des études bien menées	<b>A</b>  Preuve scientifique établie
<b>Niveau 2</b>  - Essais comparatifs randomisés de faible puissance - Études comparatives non randomisées bien menées - Études de cohorte	<b>B</b>  Présomption scientifique
<b>Niveau 3</b>  - Études cas-témoin	<b>C</b>
<b>Niveau 4</b>  - Études comparatives comportant des biais importants - Études rétrospectives - Séries de cas - Études épidémiologiques descriptives (transversale, longitudinale)	Faible niveau de preuve scientifique

La recommandation de grade A est fondée sur une preuve scientifique établie par des puissances de fort niveau, par exemple des essais comparatifs randomisés de forte puissance et sans biais majeur, méta-analyse d'essais contrôlés randomisés, analyse de décision basée sur des études bien menées.

La recommandation de grade B est fondée sur une présomption scientifique fournit par des études de niveau intermédiaire de preuves : par exemple, essai comparatif randomisé de faible puissance, étude comparative non randomisée bien menée, étude de cohorte.

La recommandation de grade C est fondée sur des études de moindre niveau de preuve : par exemple, étude cas-témoins.

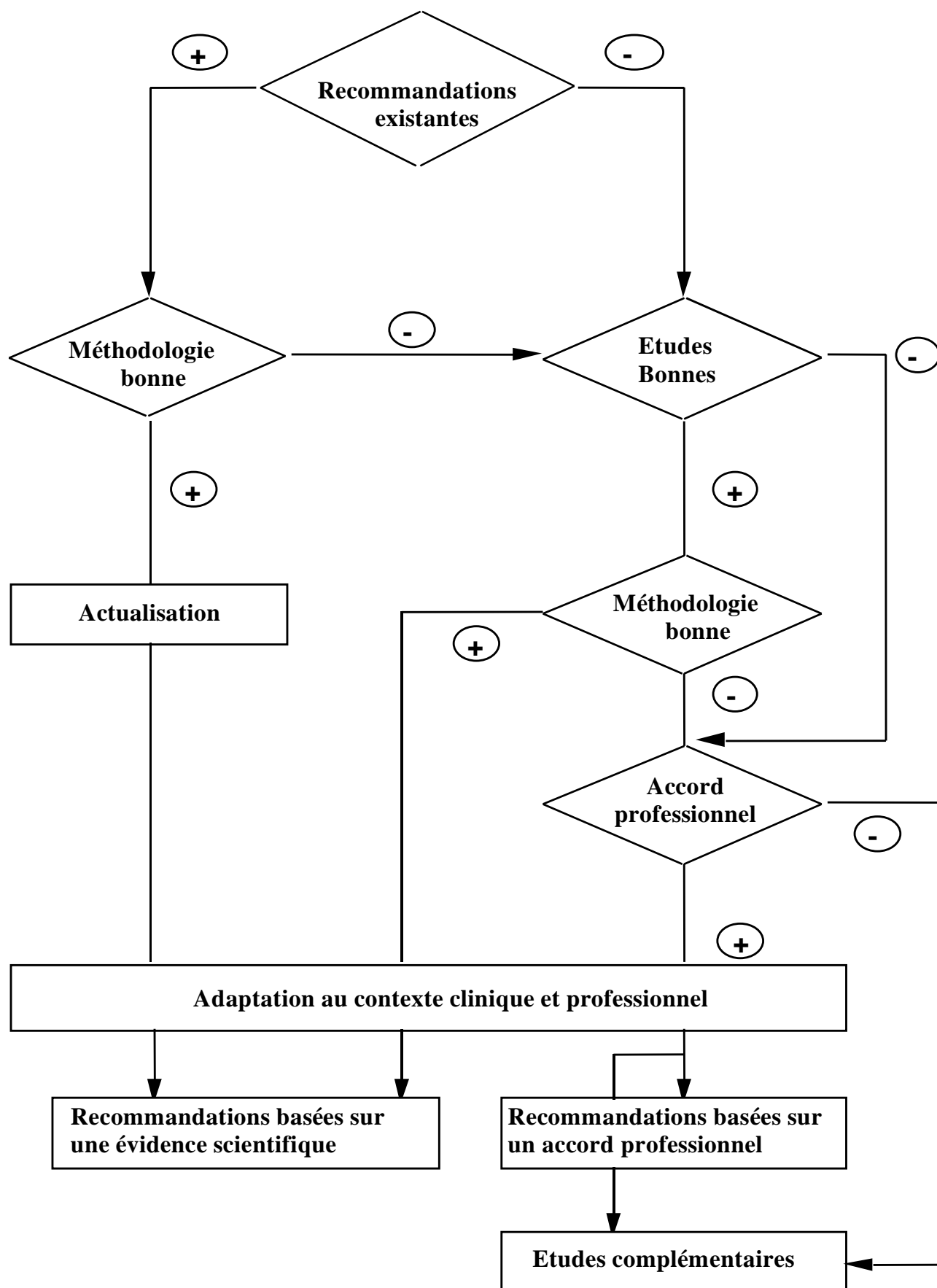
En l'absence de précision, les recommandations proposées ne correspondent qu'à un accord professionnel.

L'existence d'une évidence scientifique forte entraîne systématiquement une recommandation de grade A, quelque soit le degré d'accord professionnel.

L'analyse de la littérature permet de répondre à toutes les questions posées.

Les recommandations devront explicitement distinguer les réponses soutenues par une évidence scientifique et celles qui ne le sont pas.

La synthèse de la démarche globale d'élaboration des recommandations est résumée sur l'arbre décisionnel ci-après.



**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 améliore également un tri qui pourrait être subjectif.

Apprendre à lire est un temps essentiel de la vie médicale.

**CHAPITRE III :**

**LA SYNTHÈSE DE L'INFORMATION SCIENTIFIQUE :**

**ÉCRITURE D'UN RÉSUMÉ :**

Nous avons déjà vu toute l'importance du résumé d'un article scientifique 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 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 :
    - ü contexte (ou objectifs),
    - ü méthodes,
    - ü résultats,
    - ü conclusion.

## **II- LA CONSTRUCTION DES RESUMES :**

La construction des résumés informatifs, structurés ou non, repose sur des règles de rédaction similaire à celui du corps de l'article scientifique. La principale similitude est l'utilisation de la structure : introduction, méthodes, résultats, discussion. Chaque élément de cette structure, cependant, ne sera exposé que de manière abrégée.

Ainsi, l'introduction (ou le contexte) est limité à l'énoncé de l'objectif principal, parfois précédé d'une phrase de description du contexte.

L'énoncé des méthodes est limitée à la description succincte du schéma de l'étude, du type de la population étudiée, des mesures nécessaires à la réponse de la question principale. Cette partie « méthodes » est éventuellement complétée par une description simple des méthodes statistiques et parfois pour certaines revues, du lieu du déroulement de l'étude.

L'énoncé des résultats est aussi limité aux principales données.

Enfin, l'énoncé de la principale conclusion tient habituellement lieu de discussion.

Les autres règles de rédaction (précision, clarté, fluidité, concision) s'appliquent au résumé. La nécessité de limiter la taille, implique néanmoins que concision et précision deviennent particulièrement importantes. La concision maximale est atteinte en limitant chaque section aux éléments importants et en ne fournissant aucun détail. Le résumé se limite à l'énoncé des faits concernant les messages importants que le lecteur doit retenir.

L'absence de détails, cependant, ne doit pas être synonyme de manque de précisions. Notamment, l'énoncé du type de population doit comporter les critères d'inclusion mais se limiter aux principaux critères. Les résultats doivent aussi être précisément énoncés et quantifiés.

### **III- REGLES COMPLEMENTAIRES DE REDACTION :**

1) le résumé ne doit être rédigé qu'une fois le reste de l'article ou du rapport fini. Cela permet de garantir qu'il y a une parfaite concordance entre le contenu du résumé et celui de l'article : tout ce qui est énoncé dans le résumé doit être retrouvé quelque part dans le corps de l'article et doit être exprimé de la même manière (même définition, mêmes données ...). La rédaction tardive du résumé est plus facile, car la sélection des méthodes et des données principales aura déjà été faite, notamment dans les sections « résultats et discussions ».

2) les auteurs doivent éviter d'introduire dans le résumé des éléments susceptibles de nuire à sa clarté. Ainsi, l'utilisation d'abréviations doit être limitée. Si une abréviation permet de simplifier l'énoncé d'expression longue, plusieurs fois répétée dans le résumé (par exemple, VIH pour Virus de l'Immuno-Déficienc e Humaine), elle doit être définie selon les règles habituelles (une abréviation est introduite entre parenthèses, après l'énoncé en toutes lettres de l'expression correspondante). Le résumé ne doit inclure ni référence, ni tableau, ni figure. L'absence de référence correspond à l'obligation de limiter le contenu du résumé aux éléments directement liés à l'étude. Cela implique l'utilisation exclusive du passé (imparfait ou passé composé) pour les verbes du résumé.

**CHAPITRE IV :**  
**LECTURE CRITIQUE DES ANALYSES STATISTIQUES :**  
**METHODES ET INTERPRETATIONS**

## Plan

Introduction

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

II- Plan habituel de la section méthodologie statistique

III- Elements principaux de la lecture critique statistique

Conclusion

Annexe : pour ceux qui veulent tricher (A NE PAS RETENIR !)

## 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 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.

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 d'une étude doit être clair et une seule hypothèse doit être posée a priori,
- 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).

5 questions sont essentielles à se poser pour évaluer la qualité de la méthodologie d'une étude originale, en particulier celle d'un essai thérapeutique :

1) Cet article apporte-t-il 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 doit également apporter en termes d'importance clinique, pour le patient, pour la société, et notamment pour les décideurs.

2) Quels sujets étaient concernés par l'étude ?

Cette question peut regrouper plusieurs sous-questions :

- Comment les sujets ont-ils été recrutés ? un **recrutement exhaustif** sur une période de temps donné, ou dans un lieu donné est correcte 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 où un problème de suivi peut se poser par rapport à la compréhension de l'essai, au 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 être appliqués dans d'autres populations que la population de l'essai.

- Qui a été exclu de l'étude ? il ne s'agit pas ici de discuter de l'inverse des critères d'inclusion, mais 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.

- 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 souvent une information beaucoup plus importante que celle qui est donnée en pratique quotidienne au patient (information qui n'est d'ailleurs certainement pas suffisante) ; certains équipements peuvent être utilisés alors qu'ils ne le seront pas facilement en pratique routinière.

3) Le protocole est-il clairement et suffisamment expliqué dans l'article et les mesures effectuées suffisamment sensibles ?

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. En particulier, il existe souvent une surestimation inconsciente par les auteurs de l'originalité ou de l'importance potentielle de leur travail. Les informations données ne sont souvent pas suffisantes à propos des sujets ou de la prise en charge réalisée.

De plus, 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 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.

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

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 le plus possible 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. Ceci malheureusement est presque toujours le cas. L'absence de randomisation doit donc 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.

Enfin, les évaluations doivent être faites le plus possible **en aveugle**. Le double aveugle 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.

5) 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 était adéquat. Des données précises doivent être fournies par les auteurs, en particulier sur la médiane de suivi.

Enfin, l'exhaustivité du suivi et notamment le nombre de sujets qui ont été perdus de vue et pourquoi, doit être également précisé. Des organigrammes et des règles de rédaction d'un article rapportant un essai thérapeutique ont été mis au point par plusieurs journaux internationaux (CONSORT statement).

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.**

## **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 :

- Modalités d'analyses descriptives 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 (valeurs qui séparent 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 deviation* 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 parfois peut donner aux lecteurs une impression fautive de précision importante 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 bonnes pratiques que les auteurs aient précisé s'ils ont exprimé leurs résultats sous la forme de moyenne  $\pm$  écart type ou sous la forme de moyenne  $\pm$  erreur standard.

- 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 traiter ou non dans les analyses. Dans certaines études, des traitements spécifiques de ces valeurs manquantes peuvent être mis en œuvre. Par exemple, dans des études de cohorte, des techniques statistiques sophistiquées sont disponibles.

- 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 avoir été 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.

- Analyse des critères secondaires : les autres critères de jugement doivent être également comparés et les tests statistiques adaptés doivent être précisés.

- 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, 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 puis sélection progressive des variables pour aboutir au modèle final.

- 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.

- 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** c'est-à-dire que les auteurs se laissent 5% de risque d'erreur de conclure à une différence statistiquement significative alors qu'elle n'existe pas réellement). Ce seuil peut toutefois être amené à 0,10 pour certains résultats d'analyse multivariée. L'inclusion des variables dans le modèle initial d'une analyse multivariée peut être faite jusqu'à un seuil de 0,25.

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 de certains risques d'erreurs doit avoir été prévu et notamment 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 quels sujets seront dans un groupe et quels sujets seront dans l'autre groupe. Toute autre méthode risque de mener à des biais lors de comparaisons de résultats. 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. 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 fait appel à des tables de nombres au hasard.

La randomisation classique est équilibrée par blocs de taille variable, c'est-à-dire s'assure que les différents groupes auront un effectif identique à des moments précis au fur et à mesure des inclusions.

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 en terme d'âge, de sexe par exemple, ou des variables pronostiques importantes. Il n'y a pas lieu de réaliser des tests statistiques à ce niveau, si la randomisation a été correctement réalisée. 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. 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 indiqués avec les résultats du calcul du nombre de sujets nécessaires.

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

Parfois, le calcul de la taille des échantillons n'a pas pu être fait a priori, à cause d'une méconnaissance de l'importance de ce calcul ou d'un manque de sujets disponibles. 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 ! Toutefois, l'important est que ce test soit adapté au type de données disponibles. Les tests statistiques sont différents si l'on travaille sur uniquement des données qualitatives, sur une comparaison entre données qualitatives et quantitatives, uniquement sur des données quantitatives (tableau I). Une donnée quantitative s'exprime par une moyenne couplée éventuellement à un écart type. Une donnée qualitative s'exprime en pourcentage.

Très schématiquement, s'il s'agit de comparer des pourcentages dans deux groupes de sujets (deux variables qualitatives), le test approprié est un  $\chi^2$ . S'il s'agit de comparer une variable quantitative, c'est-à-dire deux moyennes entre deux groupes (le groupe étant une variable qualitative), 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 plusieurs moyennes (plus de deux), les tests correspondants sont une analyse de variance (test paramétrique) ou un test de Kruskal-Wallis (test non paramétrique). S'il s'agit de regarder si une variable quantitative va dans le même sens qu'une autre variable quantitative, il s'agit de corrélation, voire de régression. Les tests appropriés sont alors le calcul du coefficient de corrélation linéaire « r » (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 par exemple 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 évidemment 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. En général, les tests paramétriques sont plus puissants que les tests non paramétriques et doivent donc être utilisés si possible. Les tests non paramétriques travaillent sur le rang des données et non pas sur la valeur réelle.

Décider si les données sont distribuées de façon normale est difficile mais doit être fait puisque cela détermine quel type de test devait être réalisé.

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.

Si des auteurs ont utilisé des tests statistiques non usuels, voire obscurs, il faut que l'article justifie pourquoi cela a été réalisé et qu'ils apparaissent dans les références. Le nombre de tests statistiques possibles semblent parfois infinis. Toutefois les comparaisons statistiques habituelles peuvent être réalisées en utilisant au maximum une douzaine de tests. Si les auteurs semblent décrire 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'hypothèse nouvelles, qui devront être testées par d'autres études. Seules des raisons éthiques peuvent permettre d'arrêter prématurément un essai thérapeutique.

#### 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. 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 de montrer l'inverse.

#### 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 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 fameux 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.

A l'inverse, une valeur de « p » qui n'est pas inférieure 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.

Enfin, une valeur de « p » ne suffit souvent pas. Il faut donner les valeurs comparées et pas seulement la valeur de « p ». De plus, 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 soit positive c'est-à-dire montrer une association statistique, ou négative. 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, qu'ils soient positifs ou négatifs. 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 réalisée grâce au bon sens clinique. L'utilisation des intervalles de confiance dans les articles médicaux est encore malheureusement peu commune.

## **CONCLUSION :**

Ces quelques pages n'avaient 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 à la question de base : Quels paramètres les auteurs ont-ils voulu comparer ? et compléter cette question par : De quels types étaient les paramètres ? c'est-à-dire qualitatifs ou quantitatifs. Il existe des exemples de mauvaise utilisation ou de mauvaise interprétation des tests statistiques dans la plupart des grandes revues, même internationales. Le paragraphe qui suit donne quelques exemples qui ne doivent pas être suivis...

**ANNEXE****POUR CEUX QUI VEULENT TRICHER (A NE PAS RETENIR ET ENCORE MOINS APPLIQUER !!!) :**

- lancer toutes vos données dans un ordinateur, et rapporter comme significatives toutes relations où « p » est inférieur à 0,05 ...,
- si les différences existent initialement entre deux groupes et qu'elles favorisent le groupe où a été donné le nouveau traitement, se souvenir de ne pas ajuster sur ces différences ...
- ne pas s'intéresser à la distribution des données. Si vous le faites, vous risquez d'être obligé d'utiliser des tests non paramétriques, ce qui n'est pas très drôle,
- ignorez tous les perdus de vue et les non répondus, comme cela l'analyse concerne uniquement les sujets ayant eu l'ensemble du traitement,
- si des points extrêmes vous ennuient dans vos calculs, enlevez-les.
- mais si les extrêmes vous aident dans vos conclusions, laissez-les,
- si vos intervalles de confiance sont trop larges, voire inclus la possibilité de l'absence de différence entre les groupes, ne les écrivez pas dans l'article. Mieux encore, mentionnez les très brièvement dans le texte, mais ne les dessinez pas sur un graphe et surtout ignorez les lorsque vous donnerez vos conclusions,
- si une différence entre deux groupes devient significative quatre mois et demi après le début d'un essai alors qu'il devait durer six mois, arrêter l'essai thérapeutique et commencer à donner les résultats. Par contre, si à six mois les résultats sont proches de la significativité, continuer l'essai pour un mois supplémentaire...
- si vos résultats semblent non intéressants, demandez à l'ordinateur de revenir voir si certains sous groupes se comportent différemment. Vous trouverez peut être que votre traitement est intéressant chez toutes les femmes âgées de 52 à 61 ans et mesurant 1,57 m à 1,65 m...
- si l'analyse de vos données de la façon dont vous l'aviez planifiée ne vous donne pas les résultats escomptés, recommencer avec d'autres tests ...

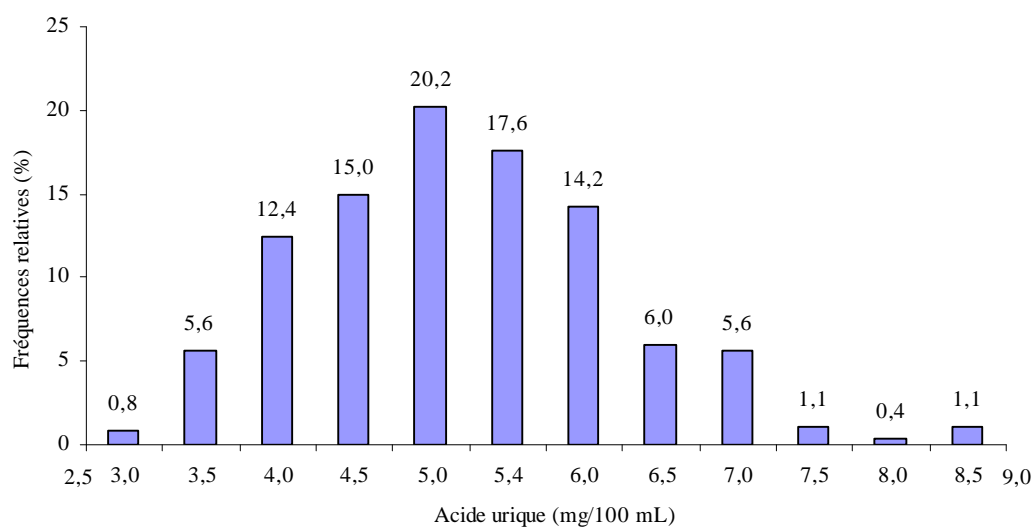
**Tableau I : Niveau de mesure des variables et choix des méthodes statistiques**  
(cas de comparaisons non appariées)

<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	$X^2$ ; test exact de Fisher
	Multinomiale	Proportions ; risques relatifs	$X^2$
	Ordinale	Proportions ; risques relatifs	$X^2$ ; test $X^2$ de tendance
	Quantitative	Moyennes ; médianes	Test $t$ ; test de Mann-Whitney
	Multiples	Régression logistique ; analyse discriminante ; modèles log-linéaires ou linéaires généralisés	
<b>Multinomiale</b>	Multiples	Régression logistique ; analyse discriminante ; modèles log-linéaires ou linéaires généralisés	
<b>Ordinale</b>	Multiples	Régression logistique ; analyse discriminante ; modèles log-linéaires ou linéaires généralisés	
<b>Quantitative</b>	Dichotomique	Moyennes ; médianes	
	Multinomiale	Moyennes ; médianes	
	Ordinale	Moyennes ; médianes	
	Quantitative	Coefficients de corrélation ou régression	
	Multiples	Régression linéaire multiple ; modèles linéaires généralisés	
<b>Temps</b>	Multiples	Modèles de survie (Kaplan-Meier, Cox, Poisson ...)	

\* Liste non exhaustive ; le choix des méthodes dépend aussi du type de question posée et des conditions d'application des tests.

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

**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