

Typologie des questions en LCA aux EDN

Arwa Ben Salah, Moustapha Dramé, Damien Jolly

L'épreuve de lecture critique d'articles

L'introduction

Les objectifs de l'étude

Le schéma de l'étude

La population de l'étude, en particulier

La maladie étudiée

Les variables étudiées

L'analyse statistique

Les résultats de l'étude

La discussion

L'épreuve de lecture critique d'articles

L'épreuve de lecture critique d'articles (LCA), telle qu'elle figure aujourd'hui aux épreuves dématérialisées nationales, est constituée de deux articles en anglais remis sous forme papier, sur chacun desquels est posé un nombre de questions à réponses multiples qui peut varier entre 13 et 17, en règle générale une quinzaine.

En fait, l'analyse des différents sujets déjà posés antérieurement et l'expérience des enseignants montrent qu'au final, à propos d'un article particulier — qu'il s'agisse d'une enquête épidémiologique, d'une évaluation diagnostique, d'un essai thérapeutique, etc. —, les questions posées sont toujours à peu près les mêmes et qu'il existe une certaine typologie limitée de questions, que nous allons essayer de développer.

En suivant l'ordre naturel de la rédaction classique d'un article, de structure IMRAD (introduction, méthode, résultats « and » discussion), on trouve, en général, des questions sur :

Glossaire en ligne



Les termes signalés par un astérisque font l'objet d'une définition dans le glossaire en ligne. Pour le consulter, connectez-vous sur <http://www.em-consulte/e-complement/478335> et suivez les instructions.

L'introduction

Quelle est la justification de l'étude ?

Que sait-on sur le sujet – Que ne sait-on pas ? Pourquoi a-t-on fait cette étude ?

Très souvent, il s'agit d'une question sur les nouveautés scientifiques ou les interrogations ayant amené à la réalisation de l'étude. C'est souvent la première question. Les réponses sont a priori écrites dans la partie « Introduction » de l'article.

Les objectifs de l'étude

Quel est l'objectif principal ?

Quels sont les objectifs secondaires ?

Ces informations se situent en fin d'introduction, après le rationnel justifiant l'étude, comme une sorte de conclusion à l'introduction, juste avant les méthodes.

La juste reconnaissance de ces objectifs*, de plus, permet très souvent de préciser comment les auteurs ont fait pour répondre à ces objectifs en spécifiant le schéma ou le design de l'étude, souvent objet d'une autre question. Très souvent, on peut déjà avoir une idée de la population* étudiée, des expositions à d'éventuels facteurs de risque, des interventions, des comparateurs et des *outcomes** (critères de jugement).

L'objet de l'article en question était-il une évaluation d'une nouvelle procédure diagnostique, d'un nouveau traitement, l'estimation d'un pronostic, la construction d'un score diagnostique ou pronostique, une étude à visée étiologique d'épidémiologie analytique, etc. ?

L'évaluation d'un nouvel examen diagnostique se fait le plus souvent en comparaison à un examen de référence, ou *gold standard*; elle implique l'estimation d'indicateurs, comme la sensibilité*, la spécificité*, les valeurs prédictives positive* et négative*, les rapports de vraisemblance* positif ou négatif, et utilise parfois les courbes ROC pour avoir une vision globale de la validité diagnostique ou pour comparer le nouvel examen au *gold standard* et, éventuellement, pour le comparer à d'autres procédures diagnostiques.

L'évaluation de nouveaux traitements, de nouvelles stratégies thérapeutiques, éventuellement d'interventions en promotion de la santé ou en éducation pour la santé repose, en règle générale, sur les techniques de l'essai clinique avec comparaison de groupes de patients au sein desquels les interventions ont été différentes. La randomisation* permet d'obtenir des groupes a priori comparables initialement. Il est important de savoir au départ si l'on est face à un essai de supériorité ou de non-infériorité*. L'analyse porte le plus souvent sur un critère de jugement mesurable reconnu dans la littérature scientifique ou, comme souvent en cancérologie, sur le délai de survenue du décès (modèles de survie).

L'évaluation de la mise en place d'un programme de dépistage évalue une intervention chez des sujets *en apparence* non malades pour permettre de faire un diagnostic plus précoce. Elle ressemble beaucoup à l'évaluation d'un examen diagnostique par les indicateurs utilisés.

L'estimation d'un pronostic s'adresse à une population de sujets malades pour laquelle on évalue des différences d'évolution. Elle fait souvent appel aux modèles de survie pour son analyse, le critère de jugement utilisé pouvant être le délai de l'apparition ou non d'une complication (souvent le décès) au cours du temps. Elle permet parfois la création et la proposition de nouveaux scores pronostiques, utilisables en pratique courante.

Les enquêtes épidémiologiques sont souvent effectuées sur des échantillons parfois représentatifs de la population générale. Elles sont soit simplement descriptives, soit analytiques (étiologiques) visant à déterminer les facteurs de risque associés à une (enquête cas-témoins*)

ou plusieurs maladies (enquête de cohorte*) chez des sujets qui, initialement, en étaient indemnes. Elles visent parfois à déterminer les effets d'une intervention sur des populations ou des groupes de sujets particuliers, reprenant ainsi des méthodes de comparaison proches de celles de l'essai comparatif.

Le schéma de l'étude

Il s'agit du schéma expérimental ou du schéma d'enquête ainsi que ses méthodes de réalisation. *S'agit-il une étude observationnelle* ou expérimentale (interventionnelle) ?*

Est-ce un essai thérapeutique, une étude de cohorte*, une étude cas-témoins*, une évaluation diagnostique, une évaluation d'un programme de dépistage, une étude pronostique, etc. ?*

En règle générale, ces informations se situent en début du paragraphe « Méthodes », parfois même, elles sont spécifiées par les auteurs dans le titre de l'article.

Les modalités de sélection de la population, des groupes comparés, la typologie de l'analyse statistique, la présentation des résultats permettent aussi d'acquiescer, si besoin, des certitudes quant à la typologie de l'étude présentée dans l'article.

S'il s'agit d'un essai thérapeutique, l'étude peut être en groupes parallèles, en cross-over, ou en plan factoriel. Elle peut aussi être en aveugle (en insu) ou en ouvert.

La population de l'étude, en particulier

Quelle est la population source ?

Quelle est la population qui a été étudiée ?

Quelle est-elle en termes de représentativité (échantillon représentatif de la population) ?

Comment a-t-elle été constituée, sélectionnée, tirée au sort ?

Quelles en sont les caractéristiques, notamment en termes démographiques ?

On peut remarquer que dans les essais cliniques, les groupes de population* de patients comparés sont sélectionnés par la définition et l'utilisation de critères d'inclusion*, de non-inclusion ou d'exclusion*. Une sélection trop forte de la population peut donc entacher les résultats d'un d'essai bien mené montrant une différence mais dont l'applicabilité en population générale, dans la « vraie vie », pourra être fortement limitée voire sans intérêt.

Quelle est la manière dont elle est représentée dans le flow chart de l'étude ?

Il est intéressant, dans le *flow chart* (diagramme de flux*), de noter les nombres de sujets initialement participants à une étude et de regarder comment se font les « pertes » de sujets au cours de l'étude; sur combien de sujets a porté l'analyse finale. Cette information est importante en termes de validité interne* de l'étude, validité qui est souvent l'objet d'une des questions finales.

Y a-t-il de gros biais de sélection de la population ?

Biais qui pourrait d'emblée entacher la validité externe* de l'étude en plus de sa validité interne*...

Les questions sur les biais* sont très fréquentes, en sachant qu'elles ne doivent pas aller plus loin que les notions de biais de sélection, de biais de classement (encore appelé biais d'information ou biais de mesure) et de phénomène de confusion.

Quelle serait la population cible ?*

Cela concerne la population visée par les résultats de l'étude, celle sur qui on pourra généraliser les résultats.

On retrouve des informations sur ce sujet à la fois dans le paragraphe « Méthodes » qui, parfois, est structuré en sous-chapitres, avec potentiellement un sous-chapitre population. L'analyse du *flow chart*, lorsqu'il existe, est une manière de retrouver rapidement des informations sur les exclusions de sujets au cours de la réalisation et du suivi de l'étude qui pourraient exposer à des biais de sélection potentiellement majeurs, entachant au final les validités interne et externe de l'étude. Le *flow chart* permet de savoir sur quels sujets a porté l'analyse finale. On retrouve également l'ensemble des éléments permettant de décrire précisément la population sur laquelle les auteurs ont travaillé, dans le tableau 1 de l'article, en règle générale : « Caractéristiques de la population » ou « Comparabilité des groupes étudiés ».

La maladie étudiée

La population étudiée est souvent soit un groupe unique de malades (cohortes descriptives, études pronostiques), soit des groupes rendus comparables de malades (essai clinique) par un tirage au sort définissant l'appartenance à l'un des groupes définis.

Il est indispensable que la définition de la maladie soit claire et précise, qu'elle corresponde, en règle générale, à ce qu'on connaît de la maladie dans la littérature scientifique.

Elle ne doit ni être floue, ni trop précise...

En sus de la définition de la maladie, les auteurs sont susceptibles de préciser des stades ou des niveaux de gravité, des tranches d'âges ou d'autres caractéristiques épidémiologiques.

Cela est très important pour la validité externe de l'étude... et peut donc être source d'inspiration pour des questions, notamment des questions récapitulatives présentées en fin de déroulement du dossier progressif.

Les variables étudiées

On s'intéresse de manière privilégiée et en premier lieu à la variable que l'on cherche à expliquer dans l'étude.

Quels sont les critères de jugement, principal ou secondaire ?

Ces critères étaient-ils simples ou composites ?

Ces critères étaient-ils des critères cliniques ou des critères intermédiaires ou de substitution ?

S'agissait-il d'un score validé (par exemple, un score de qualité de vie) ?

Très souvent, notamment dans les essais cliniques, une question est posée sur les critères de jugement* et leur pertinence, en particulier sur le critère de jugement principal défini dans un essai thérapeutique et/ou les critères de jugement secondaires ou encore sur les caractéristiques étudiées dans une étude épidémiologique.

C'est souvent une question permettant de comprendre si l'étudiant a bien cerné la problématique de l'étude. La définition de ces critères est souvent très précise avec de nombreux sous-critères qui se prêtent bien à une question à réponses multiples.

Quels sont les critères de devenir, d'évolution ?

Ce sont les variables privilégiées pour étudier l'évolution des sujets dans le suivi de l'étude. Ces variables, en général, permettent de définir des groupes de sujets qui seront comparés :

sujets décédés comparés aux survivants et aux sujets censurés dans les études de survie, sujets présentant ou non une complication dans les études pronostiques.

Quelles sont les définitions des cas et des témoins dans une étude cas-témoins* ?

Quelles ont été les variables étudiées pour expliquer la variable privilégiée (définie ci-dessus) ?

Ce sont les variables explicatives* ; en règle générale, elles sont constituées :

- de variables permettant de décrire la population ou les groupes échantillonnés en termes démographiques (âge, sexe, sex-ratio, etc.) ;
- des variables potentiellement explicatives de la variable à expliquer, reconnues comme explicatives ou dont l'influence est discutée dans la littérature scientifique ;
- des variables potentiellement explicatives de la variable à expliquer que les auteurs pensent intéressantes à étudier en fonction de leurs nouvelles hypothèses ;
- des variables statistiquement associées à la variable à expliquer en analyse bivariée.

Ces variables qui définissent souvent l'exposition à des facteurs de risque ou la présence de facteurs potentiellement pronostiques doivent disposer de définitions claires, précises et potentiellement reprendre des définitions admises dans la littérature scientifique.

Les variables étudiées sont une source d'inspiration importante pour poser des questions, notamment sur leur définition, leur qualité, leur validité, leur pertinence clinique, leur reproductibilité pour leur utilisation en pratique courante.

L'analyse statistique

Quelles sont les modalités d'analyse de l'étude ?

Quelle était l'analyse prévue ?

Quelles sont les modalités d'analyse descriptive, bivariée et multivariée ?

- *Analyse descriptive** : une seule variable étudiée à la fois, c'est ce qu'on retrouve avec la définition d'indicateurs de distribution (moyenne, écart-type, médiane, pourcentages, etc.) dans le tableau de description de la population.
- *Analyse bivariée** (parfois abusivement appelée analyse univariée) : l'étude des liens statistiques entre la variable privilégiée à expliquer et les autres variables potentiellement explicatives. On voit parfois apparaître des degrés de signification seulement lorsque des liens statistiques existent. Elle permet la sélection des variables candidates pour l'analyse multivariée. Selon la typologie de la variable, des tests différents seront appliqués (par exemple, test de Student si variables quantitatives continues, test du chi-2 si variables qualitatives, voire tests non paramétriques).
- *Analyse multivariée** : l'étude des liens entre la variable à expliquer et les variables potentiellement explicatives. Elle permet d'étudier l'existence ou non d'interactions entre les variables explicatives. Elle permet surtout de définir des estimations ajustées de paramètres statistiques comme notamment des indicateurs de risque ajustés (risque relatif, rapport de cotes [encore appelé *odds ratio*], *hazard ratio*) et de neutraliser les effets de confusion potentiels.

Quels sont les modèles utilisés ?

Selon la typologie du critère de jugement, des modèles différents seront appliqués, par exemple :

- régression linéaire* (multiple ou généralisée), si la variable à expliquer est une variable continue ;
- modèle logistique*, si la variable à expliquer est qualitative à deux (voire à trois) modalités ;
- modèle de Cox*, si la variable à expliquer est le délai de survenue ou non d'un événement binaire, etc.

Quelles sont les mesures d'association utilisées ?

Risques relatifs*, rapports de cote*, rapports d'incidence instantanée, différences de risques, etc.

Y a-t-il eu des analyses intermédiaires ? Prévues ou non ? Avec une prise en compte du risque d'inflation du risque alpha* ou non ?*

La répétition des tests statistiques entraîne une inflation non maîtrisée du risque alpha* et donc pourrait amener à trouver des différences significatives uniquement du fait du hasard. Cette volonté de maîtrise du risque de première espèce, les modalités d'analyse mises en place (diminution du seuil de signification, utilisation de tests statistiques protégés, planification des analyses intermédiaires avec des seuils de signification différents calculés) sont un des critères de validité interne de l'étude présentée.

Y a-t-il eu des analyses en sous-groupes ? Étaient-elles prévues ou bien ont-elles été effectuées a posteriori en connaissant les résultats ?

Les analyses en sous-groupes* non planifiées sont souvent considérées comme suspectes. De toutes les façons, elles entachent la généralisation potentielle des résultats, donc la validité externe*. De plus, ces analyses supplémentaires entraînent systématiquement une inflation du risque alpha*, ou risque de première espèce, liée à la répétition des tests.

L'analyse a-t-elle été faite en intention de traiter ou en per-protocole* ou en traitement reçu dans un essai thérapeutique ? Était-ce une hypothèse de biais maximum* ?*

Cette question n'a de sens que dans les essais thérapeutiques. Elle est fréquemment posée.

Quelle était la méthode de gestion des données manquantes ? Les données manquantes ont-elles été imputées ou pas ?

On doit retrouver les pistes pour répondre à ce type de questions au sein du paragraphe « Méthodes » et/ou du sous-paragraphe « Statistiques ».

Nombre de sujets nécessaires ?

Très souvent, une question est posée sur le nombre de sujets nécessaires* à inclure dans l'étude pour observer une différence cliniquement pertinente. Souvent, la question porte sur les éléments nécessaires au calcul du nombre de sujets :

- l'hypothèse uni- ou bilatérale des tests statistiques ;
- le risque alpha* (le risque de première espèce) ;
- le risque bêta* (le risque de seconde espèce) ou la puissance* ($1 - \beta$) ;
- l'estimation d'un niveau de référence du critère de jugement* : pourcentage dans le groupe témoin pour une variable qualitative, ou moyenne et sa variabilité (écart-type* ou variance*) pour une variable quantitative (bras témoin) ;
- la différence escomptée observable en termes de pourcentage ou en termes de variation de moyenne selon que le critère de jugement est qualitatif ou quantitatif ;
- et, enfin, la prise en compte des perdus de vue au cours du suivi dans une étude longitudinale.

Les résultats de l'étude

L'interprétation des résultats est la source majeure d'information pour les conclusions qu'on peut tirer d'un article médical.

Les notions de signification statistique (interprétation d'une *p-value**, d'un intervalle de confiance*), l'interprétation des différentes mesures de risque doivent être maîtrisées.

Des questions sur la taille de l'effet* (*effect-size*), sur l'effet observé statistiquement et/ou sur l'effet cliniquement pertinent peuvent être posées.

Des questions portent presque systématiquement sur la lecture des tableaux et/ou des figures et sur l'interprétation de ces résultats.

Des questions sur l'homogénéité des résultats au sein de divers groupes de population ou lors d'analyses stratifiées sont très fréquentes.

Parfois, une question sur l'interprétation du nombre de sujets à traiter* pour éviter un cas et éventuellement rapporté à son coût évitable peut être posée.

Bien évidemment, l'ensemble des réponses à ces questions se fait par une maîtrise théorique des notions utilisées — elle nécessite donc une certaine compréhension de notions méthodologiques et statistiques qui sont parfois complexes — mais aussi par une lecture attentive des résultats et des tableaux et figures de l'article.

La discussion

Au niveau de la discussion, très souvent, une question peut porter sur le fait que les objectifs de l'étude ont été atteints ou non.

Des questions sur l'interprétation globale des résultats en fonction des études similaires dans la littérature peuvent être posées.

Des questions sur la validité interne* de l'étude et sa validité externe* peuvent être posées.

Dans les essais thérapeutiques, une question sur la discussion par les auteurs de la balance bénéfice-risque peut être posée.

Il s'agit, en règle générale, de questions de synthèse sur l'article étudié, posées en général en fin de dossier progressif.

Enfin, très souvent, une question sur les niveaux de preuve apportés par l'étude, tels que définis selon la Haute Autorité de Santé arrive souvent parmi les dernières questions.

D'autres questions, assez générales, visant à tester la bonne compréhension et la vue synthétique critique que peut avoir l'étudiant, de l'article considéré, peuvent être posées, notamment :

- en termes d'applicabilité des résultats, de leur généralisation en pratique courante ;
- concernant les biais possibles qui n'auraient peut-être pas été évoqués par les auteurs et qui limiteraient la portée scientifique de l'article.