

La dentisterie fondée sur les faits : Partie VI. Évaluation critique de la littérature portant sur le diagnostic, l'étiologie et le pronostic

• Susan E. Sutherland, DDS •

S o m m a i r e

Les méthodes d'évaluation critique aident le lecteur à juger de la validité (exactitude par rapport à la vérité) et de la pertinence (utilité dans l'exercice quotidien) des résultats de la recherche. Les techniques précises d'évaluation critique peuvent varier quelque peu, selon la nature de la recherche. Dans cet article — le dernier d'une série de 6 consacrée à la dentisterie fondée sur les faits — nous proposons au lecteur judicieux des cadres pour lui permettre de poser des questions rationnelles devant servir à l'évaluation de la littérature dentaire portant sur le diagnostic, l'étiologie et le pronostic.

Mots clés MeSH : dentistry; evidence-based medicine; research design

© J Can Dent Assoc 2001; 67(10):582-5
Cet article a fait l'objet d'une révision par des pairs.

L'article précédent avait pour but d'introduire le concept de l'évaluation critique — ce processus visant à déterminer si une étude est à la fois crédible et utile pour votre patient ou votre cabinet. Des lignes directrices y ont aussi été énoncées pour aider le lecteur à faire une évaluation critique des articles sur les interventions thérapeutiques ou préventives; ces lignes directrices étaient basées sur une série de questions élaborées par le groupe de médecine fondée sur les faits de l'Université McMaster^{1,2}. Le présent article propose des outils pour aider à juger de la validité et de l'utilité des documents de recherche sur les tests de diagnostic, les rapports de causalité ou le pronostic.

Évaluer des articles sur les tests de diagnostic

Au moment d'envisager l'utilisation d'un nouveau test de diagnostic, il est important de se rappeler que les tests sont rarement exacts à 100 % et que tout test donnera des faux positifs et des faux négatifs. Les meilleurs tests sont ceux qui permettent de déceler la plupart des personnes atteintes de l'affection en question (sensibilité élevée) et d'exclure celles qui n'en souffrent pas (spécificité élevée). Les tests les plus utiles aident à poser un diagnostic précis qui permet ensuite de déterminer le traitement qui convient le mieux au patient et qui lui sera le plus bénéfique. Les questions qui suivent vous aideront à juger de la crédibilité et de l'utilité d'un article qui prétend valider un test de diagnostic³⁻⁵.

Le test de diagnostic a-t-il été comparé à un test étalon, dans le cadre d'un essai indépendant mené à l'insu?

L'étalon de référence, ou test étalon, représente ici la «vérité». Il peut s'agir d'un résultat de biopsie ou d'autopsie, d'un test sanguin reconnu ou de toute autre «preuve» qui confirme ou infirme l'existence d'un état. Cependant, aucun étalon de référence n'est parfait et, dans bien des cas, il n'en existe tout simplement pas. Le cas échéant, il incombe à l'auteur d'expliquer la série de critères ou le concept théorique qui sert de base de comparaison pour le nouveau test.

De plus, il est important que l'étalon de référence et le nouveau test de diagnostic soient interprétés par 2 chercheurs différents et qu'aucun de ces 2 chercheurs ne connaisse les résultats et les conclusions de l'autre, ni les détails du cas. Tous les dentistes ont eu l'occasion d'apprécier l'utilité d'une radiographie retrocoronaire pour le diagnostic d'une lésion interproximale présumée à l'examen clinique (ou vice versa). Alors que le fait de connaître les résultats d'un groupe ou d'une série de tests est indiqué (comme les résultats d'examen *et* les résultats radiographiques), et même utile, en exercice clinique, ces informations ont pour effet d'introduire un biais dans les recherches visant à mettre au point et à évaluer un test de diagnostic. Aussi importe-t-il, à ce stade, que les chercheurs soient tenus dans l'ignorance des résultats des autres tests et de toute autre information pertinente sur le patient.

Le test de diagnostic a-t-il été évalué chez un éventail de patients représentatif de la clientèle en exercice clinique?

Un échantillon approprié de patients est celui qui se compose de patients atteints de formes légère, modérée et grave de l'état à diagnostiquer, ainsi que d'affections qui diffèrent les unes des autres mais que l'on confond souvent. Il n'est pas nécessaire d'avoir un test pour faire la distinction entre une carie primaire et une carie profonde; par contre, un test qui aide le praticien à déterminer à quel moment intervenir durant la progression de la carie est utile. De même, les tests qui aident à distinguer la douleur odontogène de la névralgie faciale sont utiles. La valeur prédictive d'un test varie en fonction de la prévalence de l'affection ciblée. Si le test de diagnostic est validé dans un lieu d'exercice clinique hautement spécialisé tel une université ou un hôpital, là où l'affection risque d'être beaucoup plus courante qu'en cabinet, il est possible qu'il donne de «meilleurs» résultats qu'en cabinet. D'où l'importance pour les auteurs de préciser le cadre de l'étude et les critères de sélection des patients.

Est-ce que tous les patients ayant subi le nouveau test ont également été soumis au test étalon?

Dans certaines études, le test étalon n'est administré qu'aux patients ayant obtenu un résultat positif au nouveau test de diagnostic. Cependant, la validation des propriétés du nouveau test de diagnostic sera faussée et biaisée, si l'administration du test étalon ne se fait pas de façon systématique.

Le test peut-il être reproduit dans mon cabinet?

L'article doit préciser exactement comment exécuter et interpréter le test et doit aborder toutes les questions liées à la préparation du patient, aux précautions à prendre et aux effets secondaires et complications possibles.

Les résultats du test s'appliquent-ils à mon patient?

a. *Le degré d'exactitude du test chez mes patients sera-t-il comparable à celui obtenu chez les patients de l'étude?*

Si le cadre d'exercice dans l'étude est similaire et que les critères de sélection des patients sont généraux, la réponse sera «probablement».

b. *Les résultats obtenus changeront-ils mes méthodes de traitement?*

Vous devez déterminer si le nouveau test fournit vraiment de nouvelles données de diagnostic que vous n'aviez pas avant, si cette nouvelle information modifiera votre manière de traiter un problème particulier et, enfin, si ces changements seront bénéfiques pour le patient. Si la réponse à l'une ou l'autre de ces questions fondamentales est «non», alors la question de l'exactitude du test ne se pose plus.

Évaluer des articles sur l'étiologie

Il est important, dans l'exercice quotidien de la dentisterie, de comprendre les liens de causalité, et plus particulièrement de déterminer les liens qui existent avec les expositions nocives.

Il faut se demander, par exemple, quels sont les risques d'utiliser un anesthésique local à base d'adrénaline chez un patient qui souffre d'angine modérée stable et quels sont les risques de ne *pas* l'utiliser? Ou encore, l'utilisation régulière de protoxyde d'azote comme sédatif en cabinet comporte-t-elle des risques pour le fœtus d'une assistante-dentaire enceinte? Les lignes directrices qui suivent visent à vous aider à faire une évaluation critique d'un article portant sur les causes ou les conséquences⁶.

A-t-on utilisé des groupes témoins similaires?

Outre l'exposition à l'agent causal potentiel, un certain nombre d'autres facteurs «confusionnels» peuvent influencer sur les résultats d'une étude. Il est important que ces autres facteurs soient similaires dans les différents groupes témoins. Il serait contraire à l'éthique de concevoir un essai randomisé pour étudier une exposition nocive; le plus souvent, on doit alors avoir recours au deuxième protocole expérimental le plus efficace — en l'occurrence l'étude de cohortes, laquelle a pour but de former des groupes de patients exposés et non exposés, de les suivre durant une certaine période et de surveiller la manifestation de l'affection à l'étude. Si cette affection est rare ou qu'elle évolue lentement, on privilégiera l'étude cas-témoin, qui consiste à identifier des sujets atteints et des sujets témoins comparables mais non atteints. L'exposition à l'agent étiologique sera évalué d'une manière rétrospective, puis les résultats seront comparés entre les 2 groupes. Il n'y a cependant aucune garantie que les 2 groupes soient similaires, vu l'absence de randomisation dans ces 2 protocoles non expérimentaux. Qui plus est, le caractère rétrospectif de l'étude cas-témoin risque d'introduire un biais considérable. Enfin, même si les études de cas et les séries de cas suscitent la réflexion et favorisent souvent la conduite d'autres recherches, elles n'ont pas de groupe témoin et ne fournissent pas de preuve permettant d'établir les rapports de causalité.

Les expositions et les résultats ont-ils été mesurés de la même manière, chez les 2 groupes?

Des biais peuvent s'introduire durant la mesure du résultat ou de l'exposition. Lorsque, par exemple, les cliniciens savent qu'un patient a été exposé à un facteur de risque, ils ont tendance à faire preuve de plus de diligence dans l'évaluation («biais de surveillance»). On peut toutefois réduire ce biais au minimum dans les études de cohorte, en s'assurant que les cliniciens chargés de l'évaluation ne savent rien de l'exposition du patient. Il en va de même des études cas-témoin; les cliniciens peuvent ainsi être portés à poser des questions plus détaillées sur l'exposition s'ils savent que le patient a réagi positivement. De même, les patients qui réagissent positivement peuvent être plus enclins à se remémorer les événements ayant mené à l'exposition ou chercher au contraire à réduire l'importance d'une exposition (p. ex., la consommation de tabac, de drogues ou d'alcool), surtout s'ils croient que l'on portera des jugements sur eux. Pour limiter ce type de biais, les patients et les cliniciens peuvent être tenus dans l'ignorance des hypothèses de l'étude.

L'exposition s'est-elle produite avant le résultat?

Ce critère est plus facile à appliquer dans les études de cohortes que les études cas-témoin, bien que la distinction ne soit pas toujours évidente. Les patients déprimés sont-ils plus susceptibles de souffrir de névralgie faciale atypique ou la dépression est-elle une conséquence de la douleur intense constante?

Existe-t-il une relation dose-effet?

Une augmentation du niveau ou de la durée d'exposition devrait entraîner une hausse du risque d'affection ou de sa gravité.

Cette relation est-elle plausible?

A-t-on écarté les autres explications? La relation est-elle plausible sur le plan biologique et est-elle conforme aux connaissances actuelles des sciences fondamentales? Cadre-t-elle avec nos connaissances actuelles?

Puis-je appliquer les résultats à mon cabinet?

Si les caractéristiques de vos patients sont similaires à celles des sujets de l'étude, si les expositions ou les traitements décrits dans l'article se comparent à ceux de vos patients et si le protocole expérimental est rigoureux, les résultats de l'étude pourraient alors être très pertinents pour votre cabinet. Votre décision, quant à modifier ou non votre mode d'exercice, dépendra de l'ampleur des risques en cause, de la rigueur des données présentées et de la disponibilité d'une solution de rechange sûre, efficace et réaliste.

Évaluer des articles sur le pronostic

Le «pronostic» se définit comme l'issue éventuelle d'une maladie ou d'une affection et la fréquence prévue de cette issue. Les patients interrogent souvent le dentiste sur le pronostic et, plus souvent encore, sur la manière dont une intervention proposée peut modifier ce pronostic. Pensons par exemple au parent qui demande si les dents de son enfant resteront toujours droites après un traitement orthodontique. D'autres patients peuvent demander combien de temps dureront un implant et une couronne, si la leucoplasie buccale mènera à un cancer de la bouche ou si une maladie parodontale entraînera la perte de dents.

Les facteurs pronostiques font référence aux caractéristiques du patient (par exemple, ses caractéristiques démographiques ou son profil biologique), de l'affection et des autres conditions co-existantes ou conditions de comorbidité (par exemple, le diabète chez un patient souffrant de maladies parodontales), qui aident à prévoir l'issue, sans en être nécessairement la cause. La présence de ces facteurs a plutôt pour effet d'accroître ou de réduire les risques que l'issue se manifeste.

L'étude de cohortes est le meilleur protocole de recherche pour l'étude du pronostic. Lorsque l'affection à l'étude est rare ou qu'elle évolue lentement à partir du moment où l'on observe un facteur pronostique pour la première fois, on peut utiliser un protocole cas-témoin, bien que les inférences que l'on pourra tirer des conclusions d'une telle étude soient beau-

coup plus faibles. Les questions qui suivent peuvent vous aider à déterminer si les résultats d'une étude sur le pronostic sont valides et appropriés⁷.

Les patients sont-ils bien décrits et représentatifs, ainsi qu'à une étape similaire de l'évolution de leur maladie?

L'affection à l'étude doit être bien décrite et les critères visant à déterminer si le patient présente ou non cet état doivent être clairement définis. Par exemple, si les chercheurs ont réuni un groupe de patients souffrant de parodontite juvénile localisée, il importe de connaître les critères précis de diagnostic ayant servi à la sélection des patients. Comme nous voudrions sans doute connaître l'issue chez tous les patients atteints, une étude basée sur la population (pour laquelle les patients pourraient avoir été dirigés par les parodontistes de la région), au sein de laquelle seraient représentés tous les niveaux de gravité de l'affection, sera plus instructive. Si l'étude n'inclut que des cas graves, inhabituels ou réfractaires dirigés vers une clinique universitaire par des cliniciens, l'issue chez ces patients ne sera pas aussi bonne et le pronostic de la maladie paraîtra alors beaucoup plus sombre qu'il ne l'est en réalité. Il importe également que les patients soient inclus dans une cohorte à un même stade précoce de la maladie (au début), car si les patients sont choisis à différents stades de l'évolution clinique de la maladie, le pronostic sera faussé. Il se pourrait par exemple qu'il y ait déjà eu perte de dents chez des patients choisis à un stade plus avancé de la maladie. Or comme les données ne sont recueillies que sur une base prospective (pour éviter tout biais associé aux protocoles rétrospectifs), ces effets négatifs (perte de dents) ne seront pas inclus et le pronostic de la maladie paraîtra alors meilleur qu'il ne l'est en réalité. À l'inverse, la période favorable de survie ne sera pas prise en considération chez les patients inscrits à un stade tardif qui auront conservé leurs dents, et le pronostic global paraîtra alors plus sombre.

Le suivi a-t-il été assez long et complet?

Les résultats qui revêtent de l'importance sur le plan clinique, comme une lésion carieuse qui doit être restaurée ou la perte d'une dent, peuvent être longs à se manifester après qu'un facteur pronostique a été déterminé. La période de suivi doit donc être suffisamment longue pour permettre de déceler le résultat à l'étude. De plus, si le suivi est incomplet, la validité de l'étude pourrait être largement compromise. Dans les études cliniques, le suivi incomplet est souvent mal documenté et il peut compromettre la validité de manière importante. Mais comment savoir que la validité d'une étude a été compromise à cause d'un suivi incomplet? Une règle de base est de mettre sérieusement en doute les résultats de l'étude, si plus de 20 % des patients ne l'ont pas terminée. On peut également tenir compte de la proportion de patients perdus durant le suivi, par rapport à la proportion de patients qui présentent l'issue défavorable, et supposer qu'il s'agit du pire des scénarios — c'est-à-dire présumer que l'issue défavorable s'est manifestée que chez tous les patients non disponibles.

N'oublions pas que l'amélioration ou l'aggravation de l'état du patient sont 2 des principales raisons qui amènent un patient à se retirer d'une étude. Présumer le pire des scénarios est une approche plus prudente que supposer que l'issue défavorable ne s'est manifestée que chez une proportion (mais laquelle) des patients non disponibles. Si la proportion de patients ayant abandonné est appréciable mais qu'une faible proportion seulement des patients manifestent l'issue défavorable parmi ceux toujours présents, la validité de l'étude devient douteuse. Prenons par exemple une étude dans le cadre de laquelle 8 % des patients sont perdus durant le suivi. Si la proportion de patients toujours inscrits à l'étude manifestant le résultat défavorable est de 30 % et que nous présumons le pire des scénarios pour les 8 % de patients non disponibles, le taux réel de patients souffrant de l'affection pourrait atteindre 38 %. Par contre, si le taux n'est que de 2 % parmi les patients toujours inscrits à l'étude, la proportion de patients ayant abandonné (8 %) et susceptibles de tous souffrir de l'affection, prend alors beaucoup plus d'importance. Le pire scénario suggère ici que le taux réel de patients manifestant l'issue défavorable pourrait atteindre 10 % plutôt que 2 %. Vous seriez donc justifié de mettre en doute la validité de cette étude. Les chercheurs devraient comparer les caractéristiques cliniques et démographiques de tous les patients ayant abandonné à celles des patients qui ont terminé l'étude, pour voir s'il y a des différences appréciables. Il est également important de connaître les raisons de l'abandon. Par exemple, un patient tout simplement enclin à ne pas respecter ses rendez-vous peut aussi avoir tendance à ne pas se conformer, et ceci, en soi, peut constituer un important facteur de pronostic pour bon nombre d'affections⁸.

Les critères servant à juger des effets ont-ils été définis clairement et appliqués de façon objective?

Le résultat à l'étude doit être clairement défini. «L'échec d'un implant», par exemple, peut avoir de nombreuses significations en l'absence de critères bien définis. Si le jugement clinique doit intervenir dans l'évaluation du résultat, il importe que le clinicien ne connaisse aucune caractéristique du patient qui pourrait influencer l'interprétation des résultats de l'étude.

Les données ont-elles été corrigées en fonction des facteurs de pronostic extérieurs?

Les facteurs, comme l'âge et la situation socio-économique, peuvent nuire à l'évaluation du pronostic. Bien que ces facteurs ne soient pas la cause du résultat, ils peuvent être associés aux véritables facteurs pronostiques ou en être des indicateurs. L'âge, par exemple, n'est pas la cause de la polycarie évolutive; l'âge peut néanmoins être associé à la carie à cause d'autres facteurs qui y sont reliés, par exemple la modification du régime alimentaire, la xérostomie causée par les médicaments ou une dépendance fonctionnelle accrue. Les auteurs doivent préciser si les résultats de l'analyse ont été corrigés en fonction de ces autres variables.

Les patients choisis pour l'étude sont-ils similaires à mes patients? Les résultats seront-ils utiles pour choisir ou éviter des traitements ou pour conseiller les patients?

Les caractéristiques des patients doivent être décrites en détail, afin de juger de leur similarités avec vos propres patients. Le fait de connaître l'évolution clinique prévue d'une maladie peut vous aider à déterminer si vous devez intervenir, à quel moment doit se faire cette intervention et ce qu'il faut dire au patient.

Conclusion

Dans cette série, nous avons défini les principes de la dentisterie fondée sur les faits et discuté des outils nécessaires à sa mise en pratique. En formulant une question clinique bien ciblée, en exécutant une recherche documentaire efficace, en évaluant les données présentées et, s'il y a lieu, en les appliquant à vos patients, vous pourrez continuer de prodiguer des soins buccodentaires de qualité dans ce milieu qui évolue rapidement. ♦

Le Dr Sutherland est membre à temps plein du Département de la dentisterie, au Centre des sciences de la santé du Collège Sunnybrook et des femmes, à Toronto.

Écrire au : Dr Susan E. Sutherland, Département de la dentisterie, Centre des sciences de la santé du Collège Sunnybrook et des femmes, H126-2075, avenue Bayview, Toronto (Ontario) M4N 3M5. Courriel : susan.sutherland@scwhsc.on.ca.

Les opinions exprimées sont celles de l'auteure et ne reflètent pas nécessairement les positions et politiques de l'Association dentaire canadienne.

Références

1. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. 2nd ed. Boston: Little, Brown and Company; 1991.
2. Sackett D, Richardson W, Rosenberg W, Haynes R. Evidence-based medicine: how to practice and teach EBM. London: Churchill Livingstone; 1997.
3. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994; 271(5):389-91.
4. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994; 271(9):703-7.
5. Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. *BMJ* 1997; 315(7107):540-3.
6. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V. Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. *JAMA* 1994; 271(20):1615-9.
7. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. *JAMA* 1994; 272(3):234-7.
8. Haynes RB, Dantes R. Patient compliance and the conduct and interpretation of therapeutic trials. *Control Clin Trials* 1987; 8(1):12-9.