

La dentisterie fondée sur les faits : Partie V. Évaluation critique de la littérature dentaire : documents concernant la thérapie

• Susan E. Sutherland, DDS •

S o m m a i r e

La dentisterie fondée sur les faits demande de définir une question axée sur le problème d'un patient et de chercher des preuves fiables pour y répondre. Après avoir trouvé ces preuves, le dentiste doit déterminer si l'information est crédible et utile dans son cabinet en utilisant des méthodes d'évaluation critique. Le présent article — qui est le cinquième d'une série de 6 consacrés à la dentisterie fondée sur les faits — présente un cadre composé de questions conçues pour aider le lecteur à évaluer la validité et l'applicabilité des études concernant des questions de thérapie ou de prévention.

Mots clés MeSH : dental research/methods; evidence-based medicine; human research design

© J Can Dent Assoc 2001; 67(8):442-5
Cet article a fait l'objet d'une révision par des pairs.

Le besoin de renseignements valables et à jour pour répondre aux questions cliniques quotidiennes ne fait qu'augmenter alors qu'ironiquement, le temps disponible pour chercher des réponses semble diminuer. Qui plus est, un nombre surprenant de recherches publiées «sont uniquement bonnes à mettre à la poubelle». L'évaluation critique permet d'évaluer rapidement les recherches et d'éliminer celles qui ne sont pas pertinentes ou qui sont de qualité médiocre. Les 2 derniers articles de la série présentent les outils à utiliser pour faire l'évaluation critique des études selon le type de question clinique dont elles traitent. Ces concepts et ces outils, élaborés à l'Université McMaster^{2,3} par le groupe de médecine fondée sur les faits, sont maintenant utilisés dans le monde entier pour les soins dispensés dans un grand nombre de professions de la santé. Le présent article examine les techniques utilisées pour évaluer les recherches sur des questions de thérapie, tandis que le dernier présentera les techniques d'évaluation critique employées pour évaluer les documents portant sur les tests de diagnostic, les causes et la prédiction du pronostic.

Questions concernant la thérapie

Quand on envisage de recourir à une intervention thérapeutique ou préventive nouvelle, le bon sens veut qu'on recherche les meilleures preuves possibles — essais randomisés et contrôlés (ERC) et examens systématiques — avant de soumettre les patients à des traitements qui pourraient se révéler

inutiles, voire nocifs. Pour une étude clinique, l'ERC est le plan le plus solide parce que la randomisation des patients inclus dans les groupes de référence réduit le biais au minimum en assurant que les membres de chaque groupe sont aussi semblables que possible à tous les égards, sauf en ce qui concerne le traitement à l'étude. Vu l'augmentation du nombre d'ERC concernant une question précise, les lecteurs ont plus de difficulté à parcourir toute l'information disponible et à en faire la synthèse pour répondre à la question clinique. On recourt donc à des examens systématiques — parfois nommés publications «secondaires» ou recherches intégrées — qui résument, analysent et signalent les résultats combinés d'un certain nombre d'ERC. Ces examens sont exécutés avec toute la rigueur attendue des études primaires, à la seule différence que l'«unité d'analyse» est ici l'étude et non le patient.

Essais randomisés et contrôlés

Les questions suivantes vous aideront à évaluer la validité et l'importance d'une étude sur un traitement ou sur une intervention préventive^{4,5}.

La répartition des patients dans les groupes de référence a-t-elle été faite au hasard?

La première chose à faire est d'examiner si la répartition des personnes traitées a vraiment été faite au hasard. L'imputation

de chaque patient au groupe expérimental ou au groupe témoin a-t-elle été décidée entièrement au hasard, c'est-à-dire en tirant à pile ou face ou en se servant d'une autre méthode similaire? Ce type de répartition aide à garantir qu'au départ, le groupe expérimental et le groupe témoin sont constitués de personnes similaires et que les différences constatées à la fin de l'essai sont dues à l'intervention et non à un facteur de «sélection» quelconque. On recherchera des termes comme *random allocation* (répartition aléatoire), *randomly assigned* (imputation faite au hasard) ou *randomized trial* (essai randomisé) dans le titre ou dans le sommaire. S'ils ne s'y trouvent pas, on passera au titre suivant. Dans la section concernant les méthodes utilisées, on cherchera une description de la méthode suivie pour effectuer la randomisation. Si on a tiré à pile ou face ou encore si on a utilisé des enveloppes codées et scellées, des tables de nombres aléatoires ou une séquence générée par ordinateur, la randomisation a été bien faite. Toute méthode de répartition où la séquence peut avoir été devinée par quelqu'un n'est pas adéquate. Malheureusement, les méthodes de randomisation sont rarement décrites et le lecteur est laissé dans l'incertitude. Quand on lit ces documents, il est bon de se rappeler qu'une randomisation inadéquate peut exagérer de 41 % l'estimation des effets du traitement et que ces effets sont exagérés de 30 % en moyenne dans les études présentées comme randomisées, mais ne décrivant pas clairement les méthodes de randomisation utilisées⁶.

Tous les patients qui ont participé à l'essai ont-ils été pris en compte et analysés à la fin de l'étude?

Il n'est pas rare de lire qu'une étude a commencé avec un certain nombre de patients et s'est terminée avec un nombre moindre, l'auteur se contentant de dire qu'un nombre déterminé de patients «n'étaient pas disponibles aux fins de suivi». Les raisons de l'impossibilité d'effectuer le suivi peuvent être extrêmement importantes. Il arrive en effet que les patients qui ne terminent pas l'essai renseignent plus sur l'intervention que les personnes qui vont jusqu'au bout. Ces patients peuvent cesser de participer à l'essai parce qu'ils ressentent des effets secondaires (même en cas de placebo) ou parce que l'intervention les a aidés et qu'une fois leur problème ou leur trouble réglé, ils choisissent de ne pas revenir aux fins de suivi. Même lorsque l'impossibilité de faire le suivi est prise en compte et expliquée dans le document, un suivi de moins de 80 % des patients initiés est généralement considéré comme inacceptable³.

Il importe aussi que les patients soient analysés au sein du groupe auquel ils ont été affectés au hasard au début de l'étude même si, ultérieurement, ils ont changé de groupe ou ne se sont pas conformés au traitement donné au groupe expérimental ou au groupe témoin. Il s'agit du «principe de vouloir traiter» qui sert à préserver l'importante fonction de la randomisation et qui assure la répartition raisonnablement égale entre les 2 groupes des facteurs dont on ne peut avoir connaissance. Cette uniformité évite que l'intervention puisse paraître efficace alors qu'elle ne l'est pas, et elle rend les résultats de l'étude plus modérés et plus crédibles.

La répartition a-t-elle été faite à l'insu des patients, des cliniciens et du personnel affecté à l'étude?

Pour réduire au minimum l'effet placebo, les patients ne doivent pas savoir s'ils se trouvent dans le groupe actif ou dans le groupe témoin. Pour diminuer le «biais dû à la mesure», il convient aussi que le clinicien chargé d'évaluer les résultats soit tenu dans l'ignorance. Plus la répartition se fait à l'insu du personnel affecté à l'étude, plus l'essai est rigoureux.

Les groupes étaient-ils semblables au départ et ont-ils été traités de façon égale tout au long de l'étude?

La randomisation ne crée pas toujours des groupes équilibrés quant aux facteurs de pronostic connus, en particulier lorsqu'il s'agit de petites études. Les investigateurs doivent présenter des données de référence sur tous les patients inclus dans chaque groupe et, s'il existe des différences significatives, assurer le lecteur que ces différences ont été rajustées au moment de l'analyse statistique.

On appelle «co-interventions» des traitements supplémentaires autres que les traitements à l'étude, suivis par les patients ou administrés à ceux-ci. Les co-interventions posent des problèmes si elles sont réparties différemment dans le groupe expérimental et dans le groupe témoin; elles en posent beaucoup moins dans les études à double insu. Le lecteur a tout intérêt à ce que les co-interventions autorisées soient décrites dans la section concernant les méthodes et que l'étendue des co-interventions non autorisées soit documentée dans les résultats. Une fois l'étude terminée, les investigateurs peuvent évaluer le degré de l'insu en demandant aux cliniciens et aux patients de dire de quel groupe, à leur avis, le patient faisait partie, puis en comparant les réponses à la répartition réelle. Si les résultats de cette analyse montrent que le nombre de patients ou de cliniciens ayant deviné correctement la réponse est supérieur à celui auquel on aurait pu s'attendre après une répartition au hasard, c'est-à-dire si plus d'une personne sur 20, soit $p > 0,05$, a deviné correctement, les méthodes d'insu utilisées n'ont pas atteint leur but.

A-t-on évalué les résultats importants sur le plan clinique?

On considère comme «important sur le plan clinique» un résultat qui est important pour le patient. Une dent carieuse exigeant un traitement est importante pour le patient, alors que la numération des bactéries cariogènes ne l'est généralement pas. La mobilité et la perte de dents sont importantes pour le patient, alors que les mesures radiographiques de la perte osseuse ne le sont pas. Les résultats microbiologiques et radiographiques ne sont pas des résultats cliniques primaires, mais des résultats auxiliaires ou secondaires. Même s'il est important d'étudier ces résultats auxiliaires au début de la recherche sur une maladie pour aider à comprendre la progression de celle-ci, les chercheurs font souvent l'erreur de les choisir dans les essais plus décisifs parce qu'ils permettent de montrer une différence entre le groupe expérimental et le groupe témoin même si la

taille de l'échantillon est moindre et si les périodes de suivi sont plus courtes. La différence montrée peut toutefois ne pas être pertinente pour le patient. Il existe dans la littérature biomédicale de nombreux exemples de cas où de vastes essais subséquents ne réussissent pas à montrer l'efficacité d'une intervention quand ils mesurent les résultats importants sur le plan clinique, par opposition aux résultats auxiliaires⁷.

Les résultats de l'étude peuvent-ils s'appliquer à mon patient?

L'examen des critères d'inclusion et d'exclusion permet de déterminer de façon raisonnable si les résultats de l'étude peuvent être utiles pour gérer le problème du patient concerné. Si les résultats peuvent être étendus à votre patient, il importe d'examiner si les bienfaits du traitement sont supérieurs aux dommages, aux coûts supplémentaires ou aux désagréments possibles.

Examens systématiques

Les examens systématiques — également nommés «examens d'ensemble» ou «méta-analyses» si les résultats des études primaires peuvent être combinés mathématiquement — diffèrent des analyses traditionnelles présentées dans les revues ou les manuels⁸. Les examens systématiques concernent le plus souvent des questions liées à la thérapie, même s'il leur arrive aussi de porter sur d'autres questions. Des normes largement admises ont été mises au point⁹ pour l'exécution des examens systématiques liés à des questions thérapeutiques, mais il n'existe pas encore de normes et de méthodes d'évaluation critique reconnues pour les examens qui résument les résultats d'études d'observation. Les lignes directrices qui suivent vous permettront d'évaluer la validité et l'utilité des examens systématiques^{10,11} d'ERC portant sur des questions de thérapie.

Une question claire a-t-elle été posée?

La question traitée dans l'examen doit être ciblée en tenant compte de la population étudiée, de l'intervention effectuée et des résultats pris en compte. Si ces éléments clés ne figurent pas dans le titre ou dans le sommaire, vous devriez passer au titre suivant.

Les critères d'inclusion étaient-ils pertinents?

Des critères précis d'inclusion et d'exclusion liés à la population, à l'intervention, au résultat et au plan de l'étude doivent être bien définis et énoncés clairement. Cela permet au lecteur de décider si l'examen inclut les études pertinentes. Cela permet aussi de répéter l'examen et d'éviter qu'il privilégie les études soutenant un point de vue particulier.

Une recherche documentaire exhaustive a-t-elle été effectuée?

L'examen doit inclure toutes les études pertinentes et éviter d'omettre des travaux importants. Il est prouvé qu'un certain nombre d'études de qualité, solides sur le plan méthodologique, ne sont pas publiées («biais de publication») parce que leurs résultats sont négatifs¹². Les auteurs de l'examen d'ensemble se doivent de préciser leur stratégie de recherche, y compris les mots-clés et les bases de données qu'ils

ont utilisés. Idéalement, la recherche devrait inclure d'autres sources telles que des bases de données multiples, des bibliographies tirées d'articles pertinents, des travaux de conférences et des contacts personnels avec des experts.

La validité (la qualité) des études primaires a-t-elle été évaluée?

Il importe de connaître la qualité des études incluses. Si de nombreuses études sont faibles, leurs résultats combinés ne seront pas crédibles. La présence d'une critique des diverses études présentée sous forme de tableau ou d'un examen approfondi de la qualité méthodologique des études incluses est précieuse pour le lecteur.

L'évaluation des études a-t-elle été répétée par plusieurs personnes afin d'éliminer tout biais?

Pour décider de la conformité de chaque étude primaire aux critères d'inclusion, de sa validité et de la signification des données incluses, l'examineur doit porter des jugements qui sont tous sujets à des erreurs et à un biais involontaire. Pour éviter ces écueils, il est bon qu'au moins 2 des auteurs de l'examen effectuent ce processus de façon indépendante, sans connaître les décisions de l'autre personne, puis en viennent à une entente par consensus. Idéalement, le nom des auteurs des études primaires et des organismes auxquels ils sont affiliés devraient être dissimulés pendant le processus d'examen.

Les résultats sont-ils semblables d'une étude à l'autre?

Même lorsque les critères d'inclusion sont assez rigoureux, une certaine variation des résultats obtenus dans les études admissibles est inévitable. Les auteurs se doivent de présenter les traits saillants de chaque étude — patients inclus, stade ou gravité de leur maladie, intervention (dose, mode d'exécution ou moment) et méthode de mesure du résultat — et d'essayer d'expliquer la variation des résultats.

Les résultats des études ont-ils été adéquatement combinés?

Le lecteur désire savoir s'il a été raisonnable de combiner les études dans le cadre d'un examen systématique, vu qu'il n'existe jamais 2 études absolument identiques. Si des études paraissent trop dissemblables, il ne convient pas de les combiner mathématiquement. Il est possible d'effectuer un test statistique pour voir si la variation des résultats est uniquement due au hasard. Si le test indique que les résultats des études sont assez semblables pour être combinés mathématiquement, on effectue une méta-analyse. Un «comptage des votes», c'est-à-dire un dénombrement des études positives et négatives, n'est pas recommandé parce que les petites études peuvent avoir une «puissance» insuffisante, c'est-à-dire que la taille de l'échantillon utilisé peut ne pas suffire à déceler la différence d'effet du traitement sur le groupe expérimental et sur le groupe témoin. La méta-analyse a entre autres l'avantage de permettre de combiner les résultats de plusieurs études petites, mais similaires afin d'obtenir un échantillon assez grand pour déceler un effet.

Les conclusions des auteurs sont-elles corroborées par les données?

Les résultats de chaque étude doivent être rapportés de façon assez détaillée pour que le lecteur puisse juger des motifs qui sous-tendent les conclusions des examinateurs. Les conclusions sont-elles justifiées, vu la qualité méthodologique des données? Les résultats et les conclusions répondent-ils à la question posée au départ?

Les résultats aideront-ils à soigner les patients?

Comme pour toute recherche, vous devez décider si vos patients et votre lieu de travail présentent les mêmes caractéristiques que ceux sur lesquels ont porté les études incluses dans l'examen. Êtes-vous en mesure d'exécuter l'intervention dans votre cabinet et ses avantages éventuels justifient-ils les dommages ou les coûts possibles?

Conclusion

Un essai randomisé et contrôlé bien conçu est la meilleure méthodologie de recherche pour les essais cliniques. L'examen systématique est un moyen puissant de regrouper plusieurs études et de faire la synthèse de leurs résultats. Dans les 2 cas pourtant, il importe de déterminer la crédibilité de la recherche à l'aide de techniques d'évaluation critique.

Le dernier article de la série portera sur les méthodes d'évaluation critique et sur leur application aux études portant sur d'autres types de questions cliniques courantes, tels que les tests de diagnostic, l'étiologie, les causes, les dommages et le pronostic. ➤

Le Dr Sutherland est membre de la faculté 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@swchsc.on.ca.

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